

**Structural insights into diverse modes of ICAM-1 binding by
Plasmodium falciparum-infected erythrocytes**

Frank Lennartz^a, Cameron Smith^a, Alister G. Craig^b and Matthew K. Higgins^{a*}

^a Department of Biochemistry, University of Oxford, South Parks Road, OX1 3QU Oxford, UK

^b Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA, UK

* Correspondence should be addressed to: matthew.higgins@bioch.ox.ac.uk

Abstract:

A major determinant of pathogenicity in malaria caused by *Plasmodium falciparum* is the adhesion of parasite-infected erythrocytes to the vasculature or tissues of infected individuals. This occludes blood flow, leads to inflammation and increases parasitaemia by reducing spleen-mediated clearance of the parasite. This adhesion is mediated by PfEMP1, a multi-variant family of around 60 proteins per parasite genome which interact with specific host receptors. One of the most common of these receptors is intracellular adhesion molecule-1 (ICAM-1) which is bound by two distinct groups of PfEMP1, A-type and B or C (BC) - type. Here we present the first structure of a domain from a B-type PfEMP1 bound to ICAM-1, revealing a complex binding site. Comparison with the existing structure of an A-type PfEMP1 bound to ICAM-1 shows that the two complexes share a globally similar architecture. However, while the A-type PfEMP1 bind ICAM-1 through a highly conserved binding surface, the BC-type PfEMP1 use a binding site that is more diverse in sequence, similar to how PfEMP1 interact with other human receptors. We also show that A- and BC-type PfEMP1 present ICAM-1 at different angles, perhaps influencing the ability of neighbouring PfEMP1 domains to bind additional receptors. This illustrates the deep diversity of the PfEMP1 and demonstrates how variations in a single domain architecture can modulate binding to a specific ligand to control function and facilitate immune evasion.

Significance statement:

Malaria is one of the deadliest infectious diseases to affect humans, causing over two hundred million cases and hundreds of thousands of deaths annually. Its fatal symptoms occur when parasites cause infected human red blood cells to stick to human tissue surfaces, blocking blood flow and causing inflammation. This stickiness is caused by parasite PfEMP1 proteins, which interact with different human receptors, such as ICAM-1. In this paper, we demonstrate how PfEMP1 proteins bind to ICAM-1. We find that this can happen in two different but related ways, perhaps influencing which additional receptors PfEMP1 can bind. We show how the parasite can adapt to allow it to stick tightly while reducing the chance that it is detected and destroyed.

Introduction:

Despite ongoing efforts to reduce global disease burden, malaria is still one of the world's most prominent diseases, with an estimated 219 million cases each year (1). The symptoms occur as *Plasmodium* parasites divide within red blood cells of infected individuals.

Plasmodium falciparum, the cause of the deadliest form of human malaria, invades and replicates within mature erythrocytes. This intracellular habitat reduces its susceptibility to detection by the mammalian immune system, but makes it vulnerable to splenic clearance. However, the parasite also displays members of variant protein families on the surface of infected erythrocytes, including the multi-domain *Plasmodium falciparum* erythrocyte membrane proteins 1 (PfEMP1). These PfEMP1 cause infected erythrocytes to adhere to the surfaces of blood vessels and tissues, removing them from circulation and protecting the parasite within from spleen-mediated destruction (2).

PfEMP1 have evolved under conflicting selection pressures. On one hand, they have diversified into a large protein family to evade immune clearance, with around sixty antigenically distinct members encoded in each parasite genome (3, 4). Since only a single PfEMP1 is usually expressed in each infected erythrocyte, this allows for antigenic switching, enabling the population to survive as cells displaying the previous PfEMP1 variant are detected and destroyed (5). On the other hand, PfEMP1 have maintained their capacity to bind specific human cell surface proteins, often allowing them to continue to mediate adhesion to human cell and tissue surfaces even after switching to a different PfEMP1 variant. Based on their chromosomal location, the majority of PfEMP1 are categorised in three major groups; A, B and C (6, 7). Depending on the receptor to which they bind, expression of certain groups of PfEMP1 correlate with different symptoms and final outcomes of malaria episodes, with group A PfEMP1 associated with severe malaria, while the other groups tend to be linked with mild malaria, although the situation is less clear for group B PfEMP1 (8–10).

PfEMP1 bind to a variety of human ligands (11, 12), the most common of which are cluster of differentiation 36 (CD36), endothelial protein C receptor (EPCR) and intercellular adhesion molecule 1 (ICAM-1) (13–15). Group B and C (BC) PfEMP1 that contain cysteine-rich inter-domain region (CIDR) α 2-6 domains bind to CD36 and are associated with

parasites which cause mild or uncomplicated malaria (16, 17), while the A-type PfEMP1 that contain CIDR α 1 domains that bind to EPCR are associated with severe childhood malaria (14, 18–21). However, finding a correlation between expression of PfEMP1 that bind to ICAM-1 and malaria outcome has been much more challenging, with conflicting results about whether ICAM-1 binding is linked to cerebral malaria (22–25). ICAM-1 binding is mediated by a subset of Duffy binding-like (DBL) domains, the DBL β domains, which are found in either A-type or BC-type PfEMP1 (26–29). Recent studies showed that A-type PfEMP1 with an ICAM-1 binding DBL β domain also contain a neighbouring EPCR-binding CIDR α domain and that these dual-binding PfEMP1 are associated with cerebral malaria (22, 30, 31). No such correlation has been made for BC-type PfEMP1 containing a DBL β domain.

Structural studies have given significant insight into diversity and conservation of receptor binding sites in PfEMP1. Mapping sequence diversity for EPCR-binding CIDR α domains onto the structure of a CIDR α 1 domain bound to EPCR reveals that while the EPCR binding surfaces are highly diverse in sequence, they retain a specific shape and overall chemical properties (18). This allows conservation of ligand binding despite sequence diversification. This is similar to the CD36-binding CIDR α domains, where the shape and chemistry of the binding site are conserved despite extensive sequence diversity (16). In contrast, a similar analysis shows that DBL β domains from A-type ICAM-1 binding PfEMP1 are significantly less diverse. Instead, they contain a sequence motif which shows nearly total conservation in residues which directly contact ICAM-1, or that are responsible for the correct fold of the ICAM-1 binding site. However, this sequence motif is absent in DBL β domains from B- or C-type PfEMP1 that bind ICAM-1 (22). To understand how these BC-type PfEMP1 interact with ICAM-1, we have determined the structure of the DBL β domain of the B-type IT4var13 PfEMP1 in complex with the N-terminal domains of ICAM-1, revealing divergent, but similar modes of ICAM-1 binding across the PfEMP1.

Results:

ICAM-1 binding domains from BC-type and A-type PfEMP1 are evolutionary distinct

We first performed an updated phylogenetic analysis of all DBL β domains that have been tested for the ability to bind to ICAM-1 (Table S1). Since publication of similar analyses, in which the only available B- and C-type PfEMP1 sequences were from the IT4 strain of *Plasmodium falciparum* (22, 26, 28), three ICAM-1 binding DBL β domains from non-IT4 group B and C PfEMP1 have been identified (27). Inclusion of these sequences confirmed the clustering of A-type ICAM-1 binding domains observed earlier, and showed that ICAM-1 binding domains from group B and C PfEMP1 mostly form a separate, albeit less well-defined cluster (Fig. 1A). This clustering is particularly prominent in the region of the sequence which corresponds to the ICAM-1 binding site of the A-type DBL β domains, with BC-type DBL β domains lacking all of the essential features of the A-type binding site (Fig. 1B). An exception to this is the DBL β domain of the B-type PfEMP1 IT4var31, which contains many of the key features of the conserved A-type ICAM-1 binding site, but differs from the consensus sequence at crucial isoleucine and proline residues (I1078 and P1116) (Fig. 1B). This suggests that IT4var31 is either an A-type PfEMP1 that has diverged in this region, or is an intermediate between A-type and BC-type PfEMP1. With this exception, the differences between ICAM-1 binding DBL β domains indicate that A-type and BC-type PfEMP1 vary in their ICAM-1 binding sites and potentially also differ in their engagement of the receptor.

The structural basis for ICAM-1 binding by BC-type PfEMP1

To understand these differences in ICAM-1 binding and to allow comparison with the previous structure of an A-type ICAM-1 binding DBL β domain, we aimed to structurally characterise a BC-type PfEMP1 DBL β domain bound to ICAM-1. We purified a diverse set of DBL β domains, either alone or bound to the two N-terminal domains of ICAM-1 (ICAM-1^{D1D2}) and conducted crystallisation trials. To remove any flexible regions that might hinder crystallisation, we added proteases to the crystallisation drops. The DBL β domain from IT4var13 formed crystals which diffracted to 2.17 Å resolution (Table 1). We determined its structure by molecular replacement, using a truncated version of the PF11_0521 DBL β domain (Plasmo DB: PF3D7_1150400) as a search model (PDB: 5MZA) (22). A complex of the

same domain bound to ICAM-1^{D1D2} formed crystals in the presence of carboxypeptidase B and these diffracted to 3.67 Å resolution (Table 1), allowing structure determination using a known structure of ICAM-1^{D1D2} (PDB: 1IC1) (32) and a pruned version of the final structure of the IT4var13 DBLβ domain as molecular replacement search models.

The IT4var13 DBLβ domain adopts the classical DBL domain fold (33), consisting of an α-helical core decorated by extensive loops (Fig. S1A). Comparison of the unbound and the ICAM-1^{D1D2}-bound DBLβ domains showed little variation in structure upon ligand binding (root mean square deviation of 1.07 Å over backbone Cα) with variation predominantly located in loops (residues 805-809, 833-840 and 1107- 1117) (Fig. S1B).

The IT4var13 DBLβ domain interacts with domain 1 and 2 of ICAM-1 through a complex binding site, the location of which is compatible with a previous study (28), and binding is mediated by three types of interaction (Fig. 2, Fig. S1C, Table S2). A major part of the interface is mediated by hydrogen bonds, contributed by DBLβ domain side chains from a central loop (residues 973-976) and helices in the C-terminal third (subdomain 3) of the domain (residues 1098-1121) (Fig. 2A). All of these DBLβ sidechains target the backbone of ICAM-1, with the exception of glutamate 1098, which forms hydrogen bonds with arginine 49 in ICAM-1^{D1}. Secondly, a loop from the DBLβ domain (residues 1107-1117) interacts through backbone-backbone hydrogen bonds with ICAM-1^{D1} to add an antiparallel β-strand to the A'GFC β-sheet of ICAM-1^{D1} (Fig. 2B). Such β-strand augmentation is a common motif in protein-protein interactions (34), but has not been observed in other PfEMP1-receptor interactions. In the unbound DBLβ domain, this loop adopts a different conformation which is stabilised through crystal contacts (Fig. 1B, Fig. S1D), indicating that it is flexible and becomes stabilised upon interaction with ICAM-1. This is analogous to a different intrinsically disordered loop that becomes ordered to form part of the ICAM-1 binding site of A-type PfEMP1 (22). Finally, the interface is stabilised by a cluster of hydrophobic residues found on an elongated helix of the DBLβ domain and two subsequent loops (residues 1102-1139) (Fig. 2C). This cluster complements a hydrophobic patch on ICAM-1^{D1} which has previously been identified as important for ICAM-1 binding by B-type PfEMP1 (27, 35, 36).

Limited sequence conservation of the BC-type ICAM-1 binding site

We next assessed the role of different residues for binding of the IT4var13 DBL β domain to ICAM-1 by surface plasmon resonance (SPR). Since the interaction involves a flexible loop that becomes ordered upon ICAM-1 binding, we first used an SPR based assay to determine if the binding event involves a time-dependent conformational rearrangement. In such a case, the amount of stable complex formed can depend on the association time, with longer association times required to allow the second step of a two-step binding process to take place, resulting in larger quantities of stable complex and slower dissociation rates (37). Indeed, the binding of IT4var13 DBL β to ICAM-1 showed such behaviour, with the dissociation rate decreasing with an increase in association time, suggesting reordering of components of the binding site (Fig. S2). For this reason, we fitted subsequent SPR data to a two-state binding model which better describes such interactions involving a conformational change, determining an affinity of 1.22nM for IT4var13 DBL β to ICAM-1 (Fig. 3, Table S3).

We next analyzed through mutagenesis which features of the binding site are essential for interaction with ICAM-1, aiming to identify conserved markers of ICAM-1 binding among BC-type DBL β domains. We mutated each of the sidechains in the IT4var13 DBL β domain that directly contact ICAM-1 through hydrogen bonds or non-polar interactions and tested them for binding to ICAM-1^{D1D5}-Fc by SPR (Fig. 3). We found that only one (Q1103) of the five sidechains that hydrogen bond with ICAM-1 plays a major role in binding, with the Q1103A mutation causing a 200-fold decrease in affinity (Fig 3, Table S3). In contrast, mutation of three of the four hydrophobic residues that contact ICAM-1 (L1102, I1106 and F1113) reduce the interaction affinity by more than 200-fold (Fig. 3, Table S3), highlighting the importance of this hydrophobic patch for ICAM-1 binding. None of these mutations disrupted the structure of the DBL β domain, as determined by CD spectroscopy and thermal melt experiments (Fig. S3). We also mutated three glycines that form part of the loop that interacts with ICAM-1 by β -strand augmentation to test whether they confer flexibility that allows the loop to adopt its bound conformation, but these mutations did not affect the interaction (Fig. 3, Table S3).

To determine whether these binding properties are conserved, we analysed the degree of conservation of the interacting residues among the known ICAM-1 binding DBL β domains from other BC-type PfEMP1, excluding the outlier IT4var31. Sequence alignment revealed limited conservation in sequence and chemical properties among these residues, with hydrophobic residues exchanged for polar residues and vice versa (Fig. 4A). The only notable exception is phenylalanine, F1113, which forms part of the hydrophobic patch that binds to ICAM-1 and is an aromatic or hydrophobic residue in all domains analysed. Despite the lack of sequence identity, we noted that all ICAM-1 binding DBL β domains display surface-exposed hydrophobic and aromatic residues along the top third of the elongated α -helix that forms part of the binding site, albeit in positions that vary within a range of four helical turns (Fig. 4A and 4B). To test whether these residues are equivalent in function to the hydrophobic patch in IT4var13, we introduced mutations into the DBL β domain of the C-type PfEMP1 J1a (27) and tested their effect on binding to ICAM-1 by SPR. With the exception of L1093D, these mutations reduced binding and affinity by 2 to 20-fold, which was not due to the disruption of the overall fold of the domain, as verified by CD spectroscopy and thermal melt experiments (Fig. 4C, Fig. S4 and Table S4). Of these residues, Y1115 in J1a is equivalent in position to F1113 in IT4var13 and performs an equivalently important role in binding. Y1097 and L1096 in J1a form a hydrophobic patch which appears functionally equivalent to the patch formed by residues L1102 and I1106 in IT4var13, despite being found in a different region of the sequence and a different location on the helix. This use of similar binding features, contributed by different regions of the domains illustrates significant plasticity in how the BC-type DBL β bind to ICAM-1.

The flexible loop that binds ICAM-1 by β -strand augmentation in IT4var13 DBL β is also present in other BC type PfEMP1 (Fig. 4A). To test whether these use a similar mechanism of β -strand addition to bind ICAM-1 we once again assessed the dependence of dissociation rate on association time (Fig. S2). For each of the tested DBL β domains from BC-type PfEMP1 (IT4var13, J1a and Bc12) the specific dissociation rate after binding to ICAM-1 decreases with an increase in association time, indicating that the interaction involves reordering of components of the binding site (Fig. S2). We observe the same effect for DBL β

domains from A-type PfEMP1 (Pf11_0521 and PFD1235w), in which a different flexible loop becomes ordered during ICAM-1 binding (22), but not for CIDR α domains (HB3var3 and IT4var20) that bind to their receptor EPCR without any apparent structural rearrangement (18).

While the flexible loop varies significantly in length between B- and C-type ICAM-1 binding DBL β domains, most variants contain a glycine that follows the conserved hydrophobic or aromatic residue (Fig. 4A). While these glycine residues have no significance for ICAM-1 binding in IT4var13 DBL β (Fig. 3) they might be necessary for longer variants of the loop to adopt the ICAM-1 bound conformation. To test this, we mutated this glycine in the J1a DBL β and found a 7-fold reduced binding level and 1.5-fold reduced affinity, highlighting the importance of flexibility of this elongated loop (Fig. 4C).

In summary, the ICAM-1 binding site of BC-type PfEMP1 is defined by the overall chemical nature of key elements such as a surface-exposed hydrophobic patch and a flexible loop that becomes ordered upon ICAM-1 binding. However, the position, sequence and length of these elements vary significantly, in stark contrast to the highly conserved binding site of A-type PfEMP1.

Comparison of ICAM-1 binding by DBL β domains from A- and BC-type PfEMP1

Phylogenetic analyses of ICAM-1 binding DBL β domains (Fig. 1) and the significant variation in sequence conservation between the binding sites from A-type and BC-type PfEMP1 (Fig. 5A) suggests that two variants of ICAM-1 binding sites have evolved and exist in parallel, a property so far unique among PfEMP1-receptor interactions. We therefore asked what the specific differences between these two types of binding sites are and whether these would have any functional consequences for ICAM-1 binding.

The A-type PfEMP1 PF11_0521 and the B-type PfEMP1 IT4var13 have ICAM-1 binding sites in similar locations. However, the B-type binding site is more compact and the overall structure of the DBL β domain in this region is more similar to that of previously characterised non-ICAM-1 binding domains (Fig. 5B and Fig. S5) (38). In contrast, the A-type

binding site protrudes from this framework through an elongated α -helix and a loop that, in the ICAM-1 bound conformation, project away from the domain. Interestingly, this also results in a different binding angle for ICAM-1 for these two domain types. An overlay of the two structures, based on the DBL β domain, shows that while domain 1 of ICAM-1 is in a similar position in both structures, the protruding A-type ICAM-1 binding site fixes the C-terminus of domain 2 at an angle that is $\sim 30^\circ$ different from the position that it adopts, relative to the DBL β domain, when bound to the B-type IT4var13 (Fig. 5C).

To test whether this difference in angle is a general property of A- and BC-type PfEMP1, we produced two additional ICAM-1 binding DBL β domains from each of the A-type (PFD1235w and KF984156) and BC-type (Bc12a and J1a) PfEMP1. We then collected solution small angle x-ray scattering (SAXS) data for all six domains, either alone or in complex with ICAM-1^{D1D2} (Table S5) and used these data to guide molecular docking experiments, with the crystal structures of ICAM-1^{D1D2}, the PF11_0521 and IT4var13 DBL β domain structures and homology models of the other domains as input for HADDOCK (39, 40). The models of the DBL β ICAM-1^{D1D2} complexes obtained after SAXS filtering (Fig. S6A) show that this method reproduces the complexes observed in the crystal structures with high precision (root mean square deviation over backbone C α of 0.83 Å for PF11_0521 DBL β -ICAM^{D1D2} and 1.47 Å for IT4var13 DBL β -ICAM^{D1D2}) (Fig. S6B). More importantly, it shows that the resulting models fit into two distinct classes with respect to the ICAM-1 angle, with all of the A-type DBL β domains in the same class as PF11_0521 while the BC-type DBL β domains group with IT4var13 (Fig. 5D and E). *Ab initio* envelopes calculated from the SAXS data further support this observation and indicate that the differences are not due to significantly different shapes of the DBL β domains themselves (Fig. S7). Taken together, these data show that the structural variations between the two types of ICAM-1 binding site result in different binding angles for this receptor and that this difference is consistent between A-type and BC-type PfEMP1.

Discussion:

The PfEMP1 DBL β domains that bind to ICAM-1 are unusual in that they segregate into two evolutionarily distinct clusters, the A- and BC-types (Fig. 1) (22, 26). In this study we compare our structure of a BC-type ICAM-1 binding DBL β domain with the previously determined structure of an A-type domain (22), revealing that both contain a globally similar ICAM-1 binding site, in a similar location. However, these two domains differ in the conformation of the ICAM-1 binding region, the degree of sequence conservation of the binding site and the angle at which ICAM-1 is bound relative to the DBL β domain. This raises the question of why the parasite uses two similar and yet distinct types of binding sites to interact with the same receptor.

The BC-type DBL β domains and other PfEMP1 domains share the overall chemical nature of their interactions with their ligands. The core of their binding site is a hydrophobic cluster, which is complemented by hydrogen bonds contributed by both side chains and backbone groups. These features are similar to those by which DBL β domains from A-type PfEMP1 interact with ICAM-1 (22). A similar picture is also seen in CIDR α domains that interact with EPCR and CD36 (16, 18), which bind their receptors through hydrophobic or aromatic residues that either protrude from the domain or form a cavity that accepts hydrophobic residues from the receptor. Therefore, in all of the structurally characterised PfEMP1-receptor interactions, a hydrophobic patch lies at the heart of the binding site and mutation of these hydrophobic residues has a major effect on the overall affinity, with changes in both association and dissociation rates (Tables S3 and S4) (16, 18, 22). In addition, in all four cases, adjacent residues mediate hydrogen bonds that stabilise the interaction and generate increased surface complementarity with the receptor. This conservation of a hydrophobic core, together with larger complementary hydrophilic surfaces, allows the formation of stable PfEMP1-receptor complexes, enabling infected red blood cells to withstand the forces of blood flow during cytoadhesion and to evade splenic clearance.

Despite these shared features, there are significant differences in the degree of sequence conservation between ICAM-1 binding sites of the A- and BC-type DBL β domains. The

residues through which BC-type PfEMP1 bind to ICAM-1 are highly variable in sequence, and the relative positions of interacting residues and the length of the loop that forms a β -sheet addition upon ICAM-1 binding differ between domain variants. This is conceptually similar to CIDR α domains that bind EPCR or CD36 (16, 18), which also retain the overall shape and the chemistry of their binding sites, while at the same time diversifying in sequence. In marked contrast, in the A-type PfEMP1 that bind ICAM-1 all residues that are critical for direct interaction with the receptor, or the positioning of these side chains, are absolutely conserved (22).

So, which is more common, a binding site with a high level of sequence variation or one in which interacting residues are conserved? In the context of the pressure to diversify to maintain the capacity for immune evasion, theory would predict the former. Indeed, PfEMP1 containing the highly variable EPCR, CD36 and ICAM-1 binding sites are abundant in the PfEMP1 repertoire of parasite genomes from reference strains or field isolates. For example, in the 3D7 genome, 84% of PfEMP1 contain domains that bind CD36 and 11% contain EPCR-binders while, in the IT4 genome, 11% of PfEMP1 contain BC-type ICAM-1 binding DBL β domains (41). In contrast, A-type ICAM-1 binding sites occur at an average frequency of one per genome (\sim 1.3% of PfEMP1) (42). Therefore, a highly conserved binding site is the exception rather than the rule.

This raises the question of why the parasite has evolved and maintained two similar and yet distinct classes of binding sites to interact with the same receptor. In particular, the highly conserved A-type ICAM-1 binding site would appear disadvantageous under immune pressure. Indeed, antibodies that block EPCR-binding by CIDR α 1 domains have very limited cross-inhibitory activity against other variants of this domain (43). In contrast, antibodies that target the A-type binding site can broadly cross-inhibit ICAM-1 binding, and children in malaria-endemic regions rapidly acquire such antibodies (22, 44–46). A potential explanation for retention of the A-type binding site, despite it being the target of such a cross-inhibitory immune response, is that it confers a specific advantage compared to the more variable BC-type binding site. Indeed, some studies suggest that A-type PfEMP1 are preferentially expressed in individuals with limited immunity and that antibodies against

these PfEMP1 are acquired first (47, 48), indicating that they can confer an early survival advantage for the parasite.

A possible reason why parasites have evolved both A- and BC-type ICAM-1 binding domains is that, while these domains bind to ICAM-1 equally well (22, 27, 30, 49–51), they may differently influence the ability of neighbouring domains to bind additional receptors. Previous studies have shown that all PfEMP1 that contain a DBL β domain with an A-type ICAM-1 binding site also contain an adjacent CIDR α 1 domain with an EPCR-binding site, and that these PfEMP1 can bind simultaneously to both receptors (22, 30). Such dual-binding is not limited to A-type PfEMP1, since some BC-type PfEMP1 containing an ICAM-1 binding domain also have a CD36-binding domain (8, 15, 29) and ICAM-1 and CD36 have been shown to cooperate to enhance the binding of infected erythrocytes to microvascular cells (15, 52). However, in A-type PfEMP1, dual-binding to ICAM-1 and EPCR have been shown to specifically enhance binding of infected erythrocytes to endothelial cells under physiologically higher shear stresses (22, 50), which offers a distinct advantage to the parasite, especially for cytoadherence to brain endothelial cells where CD36 is absent.

The ability of a PfEMP1 to bind simultaneously to two membrane bound receptors will depend on its architecture and on how it presents binding site. The majority of PfEMP1 are thought to be rigid, elongated proteins (49, 53), constraining the conformation in which they can interact with two membrane bound receptors simultaneously. As the distance between the host cell membrane and the PfEMP1 binding site is ~ 66 Å for CD36 and ~ 32 Å for EPCR, it is tempting to speculate that the A-type and BC-type DBL β domains have evolved to bind ICAM-1 at different angles in order to allow the neighbouring CIDR α domain better to interact with either EPCR or CD36 at different heights from the host endothelial membrane. Future studies and structural insights into full-length PfEMP1 will be needed to test this hypothesis.

In summary, our characterisation of the ICAM-1 binding site of the BC type PfEMP1 supports the view that the majority of PfEMP1 utilise a binding site that is highly variable in sequence but conserved in shape and chemistry, allowing these proteins to retain receptor binding

capacity while evading immune detection. This also highlights the unusual conservation of the A-type ICAM-1 binding site, reinforcing the view that it is a suitable target for the development of anti-disease malaria vaccines.

Materials and Methods:

Protein expression and purification

The constructs for the PF11_0521, PFD1235w, KF984156, BC12a and J1a DBL β domains were described previously (22, 26, 27). Constructs comprised the following amino acids of the respective PfEMP1: 728-1214 of PF11_0521, 739-1221 of PFD1235w, 546-1036 of KF984156, 813-1273 of Bc12a, 739-1195 of J1a and 733-1202 of IT4var13. All constructs had an N-terminal hexa-histidine tag, followed by a Tobacco Etch Virus (TEV) protease cleavage site in the pET15b expression vector. They were expressed in Shuffle 3030 *Escherichia coli* (New England Biolabs) at 25°C for 16h. The DBL β domains were purified by affinity chromatography using nickel-nitrilotriacetic acid agarose (Ni-NTA, Qiagen), followed by size exclusion chromatography using a HiLoad Superdex 75 16/60 column (GE Healthcare). Mutants of IT4var13 and J1a DBL β were made using the Quikchange site-directed mutagenesis protocol (Agilent Technologies) and expressed and purified as described for wild type protein.

Constructs for ICAM-1^{D1D5}-Fc and His-tagged ICAM-1^{D1D2} were described previously (49, 54). ICAM-1^{D1D5}-Fc comprises amino acids 1 to 480 of human ICAM-1, fused to the Fc part of human IgG in a mammalian expression vector (54). ICAM-1^{D1D2} comprises amino acids 28 to 212 of human ICAM-1, fused to an C-terminal hex-histidine tag in the pHLsec expression vector (55). All constructs were transiently expressed as secreted proteins in HEK293F cells (Life Technologies). For ICAM-1^{D1D2}, kifunensine (Cayman Chemical) was added to the medium during transfection to a final concentration of 1.5 μ M. Seven days after transfection, the cell culture supernatants were harvested and sterile-filtered. ICAM^{D1D2} was then purified by affinity chromatography using Ni-NTA affinity (Qiagen), and ICAM-1^{D1D5}-Fc was purified by affinity chromatography using a HiTrap Protein A HP column (GE Healthcare). Both proteins were further purified by size exclusion chromatography using a

HiLoad Superdex 75 16/60 column (GE Healthcare). The HB3var and IT4var20 CIDR α domains and EPCR were produced as described previously (18).

Crystallisation

For crystallisation, Ni-NTA purified IT4var13 DBL β was further purified by size exclusion chromatography using a HiLoad Superdex 75 16/60 column (GE Healthcare) into 10mM HEPES, 150mM NaCl, pH 7.2. For complex formation, IT4var13 DBL β was mixed with a 1.5-fold molar excess of ICAM-1^{D1D2} and purified as described for IT4var13 DBL β . Fractions containing pure IT4var13 DBL β or IT4var13 DBL β -ICAM^{D1D2} complex were pooled and concentrated to 23 mg/ml. Crystals were grown by vapour diffusion in sitting drops by mixing 100nl of protein solution with 100nl of well solution. To remove flexible loops, Carboxypeptidase B (Sigma Aldrich), Chymotrypsin (Sigma Aldrich) or Endoproteinase GluC (New England Biolabs) were added to the crystallisation drop in a 100:1 (protein:protease) molar ratio. Crystals of IT4var13 DBL β grew at 291 K in the presence or absence of protease in conditions from the JCSG+ screen (Molecular Dimensions) containing 0.2M Magnesium chloride hexahydrate, 0.1M BIS-Tris pH 5.5 and 25% (w/v) PEG 3350. For cryo-protection, crystals were transferred into well solution containing 25% glycerol and were flash-frozen in liquid nitrogen. Crystals of the IT4var13 DBL β -ICAM^{D1D2} complex grew only in the presence of Carboxypeptidase B at 291 K in conditions from the ProPlex screen (Molecular Dimensions) containing 0.1 M Tris pH 8.0 and 25% (v/v) PEG 350 MME. The crystals were cryo-protected by transfer into well solution containing 0.1 M Tris pH 8.0 and 32% (v/v) PEG 350 MME and were flash-frozen in liquid nitrogen.

Data collection, phasing and refinement

Data for crystals of IT4var13 DBL β were collected at beamline IO4-1 (Diamond Light Source, UK) using x-rays at a wavelength of 0.92 Å and a Pilatus 6M-F detector (Dectris, Baden-Daettwil, Switzerland). Data for crystals of the IT4var13 DBL β -ICAM^{D1D2} complex were collected at beamline IO3 (Diamond Light Source, UK) using x-rays at a wavelength of 1.0 Å and an Eiger2 X 16M detector (Dectris, Baden-Daettwil, Switzerland). Both datasets were processed using the XIA2/DIALS pipeline (56) in the CCP4i2software suite (57, 58) for indexing and scaling.

The structure of IT4var13 DBL β was solved by molecular replacement with Phaser (59) using a search model consisting of a poly-alanine model of the PF11_0521 DBL β domain (PDB: 5MZA) (22) in which helices in subdomain 3 were shortened and all loops were removed. Molecular replacement found one copy of the search model in the asymmetric unit and the remaining parts of IT4var13 DBL β were built through iterative cycles of model building in Coot (60) and refinement in Phenix (61).

The structure of the IT4var13 DBL β -ICAM^{D1D2} complex was solved by molecular replacement with Phaser using pruned poly-alanine models of the IT4var13 DBL β domain and ICAM-1^{D1D2} (PDB: 1IC1) (32) as search models. To allow for flexibility between domain 1 and domain 2 of ICAM-1^{D1D2}, the domains were used as separate search models. Molecular replacement found one copy of the complex in the asymmetric unit and the remaining parts were built through iterative cycles of model building in Coot and refinement in Phenix. The structures were refined to final Ramachandran statistics of 95.6% residues in the favoured regions, 4.4% in the allowed regions and no residues in the disallowed regions for IT4var13 DBL β and 95.0% residues in the favoured regions, 5.0% in the allowed regions and no residues in the disallowed regions for the IT4var13 DBL β -ICAM^{D1D2} complex. The coordinate and structure factor data are deposited in the protein data bank (PDB) under the accession codes 6S8T (IT4var13 DBL β) and 6S8U (IT4var13 DBL β -ICAM^{D1D2}). All figures showing structures were prepared with PyMol (Schroedinger LLC).

Surface plasmon resonance spectroscopy (SPR)

SPR experiments were conducted on a Biacore T200 instrument (GE Healthcare). All proteins were purified by size exclusion chromatography and only fractions with >95% purity were used for SPR. To test binding of DBL β wild type and mutant proteins to ICAM-1, ICAM-1^{D1D5}-Fc was immobilised to 850 response units (RU) for IT4var13 DBL β proteins or 490 RU for J1a DBL β proteins, on a CM5 chip (GE Healthcare) pre-coupled with Protein A (Sigma Aldrich). All DBL β proteins were buffer-exchanged into 20mM HEPES, 300mM NaCl, 0.005% Tween-20, pH 7.2 and concentration series (0.9nM-250nM for IT4var13 DBL β , 0.9nM-500nM for J1a DBL β) were injected over the chip at 30 μ l/min, with 240s association

time and 600s (for IT4var13 DBL β) or 720s (for J1a DBL β) dissociation time. After each run, the chip was regenerated by injecting 10mM Glycine pH 1.7 for 120s at 10 μ l/min.

To test the effect of association time on dissociation rate, ICAM-1^{D1D5}-Fc was immobilised to 490 RU on a CM5 as described above, and biotinylated EPCR was immobilised to 340 RU on a CAP chip using the Biotin Capture Kit (GE Healthcare). DBL β or CIDR α domains were then injected over the chip at a fixed concentration of 3 μ M for 60s, 120s, 180s or 240s with a dissociation time of 300s. After each run, the chip was regenerated by injecting CAPture Kit regeneration solution (GE Healthcare) (for CIDR α binding to EPCR) or 10mM Glycine pH 1.7 (for DBL β binding to ICAM-1) for 120s at 10 μ l/min.

All data were analysed using the BIAevaluation software 2.0.3 (GE Healthcare). Kinetic values were determined by globally fitting the curves into a two-state reaction model. All SPR experiments were performed in duplicates, and curves shown are representatives of these measurements.

Size exclusion chromatography coupled small angle x-ray scattering (SEC-SAXS)

All SEC-SAXS experiments were carried out at the B21 beamline (Diamond Light Source, UK), using x-rays at a wavelength of 0.99 Å and an Eiger 4M detector (Dectris, Baden-Daettwil, Switzerland) with a detector-sample distance of 4.014 m. For data collection, samples were concentrated and injected at 20°C over a Superdex 200 Increase 3.2/300 column equilibrated with 20mM HEPES, 150mM NaCl, pH 7.2, with 2s exposure for each frame. The data were processed using the ScÅtter (62) and ATSAS (63) software suites, and buffer frames were averaged and subtracted from averaged frames corresponding to peak fractions. The radius of gyration (R_g) was calculated by Guinier analysis using AutoRg in PRIMUS (64). The distance distribution function $P(r)$ and the maximum particle diameter D_{max} were determined using GNOM (65). To generate volumetric representations of envelopes, 20 *ab initio* bead models were generated using DAMMIF (66). These models were then averaged with DAMAVER (67) followed by refinement against the original data using DAMMIN (68). The resulting bead models were then used to calculate envelopes with Situs. Crystal structures of the PF11_0521 and IT4var13 DBL β domains, either alone or in

complex with ICAM-1^{D1D2}, as well as models of the other DBL β domains, alone or in complex with ICAM-1^{D1D2}, were fitted into the envelopes using Chimera (69). All figures showing envelopes were made with PyMol (Schroedinger LLC).

Circular dichroism (CD) spectroscopy

All CD spectra were recorded on a J-815 Spectropolarimeter (Jasco), connected to a Peltier temperature control unit. All samples were dialysed against 100mM sodium phosphate buffer, 200mM NaF, pH 7.2 and adjusted to 0.3 mg/ml. CD spectra were recorded at 20°C using a cell with a path length of 1mM at wavelengths between 195 nm and 250 nm and spectra recorded for buffer were subtracted from these measurements. For thermal melt experiments, spectra were recorded between 200 nm and 250 nm, and between each measurement, the temperature was increased by 0.5°C increments.

Molecular modelling and docking experiments

Homology models of DBL β domains used for molecular docking experiments were generated with SwissModel (70), using structures of the ICAM-1 bound form of PF11_0521 DBL β (for KF984156 DBL β and PFD1235w DBL β) or IT4var13 DBL β (for Bc12a DBL β and J1a DBL β) as templates. To assemble complexes of DBL β domain and ICAM-1^{D1D2} by molecular docking, HADDOCK (39) was used. For this, the structures or models of the DBL β domains and ICAM-1^{D1D2} were used as input for the HADDOCK web server. DBL β residues identified to be critical for interaction of A-type PfEMP1 (22) or BC-type PfEMP1 (this study) with ICAM-1, and ICAM-1 residues known to be important for the PfEMP1-ICAM interaction (22, 27), were defined as active residues, while passive residues were chosen automatically. Other parameters were as default. Models from each docking experiment were filtered against solution scattering data from SEC-SAXS, as described by Karaca et al. (40) and the top scoring model was selected as the final docked complex.

Phylogenetic analysis

For phylogenetic analysis of ICAM-1 binding and non-binding DBL β domains, 59 sequences of DBL β domains with known ICAM-1 binding phenotype (Table S1) were aligned using MUSCLE (71). The evolutionary history was inferred using the Maximum Likelihood method

based on the Whelan And Goldman + Freq. model (72) using Mega 7 (73). The tree with the highest log likelihood (-25861.52) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories; +G, parameter = 1.1732). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 14.79% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were 360 positions in the final dataset. The tree was visualised using FigTree Version 1.4.3. A sequence logo for residues involved in ICAM-1 binding by A-type or BC-type PfEMP1 was generated using WebLogo 3 (74), based on 145 protein sequences containing the A-type ICAM-1 binding motif (22) or 10 sequences of BC-type PfEMP1 known to bind ICAM-1 (27, 29, 75).

Data availability

Data for the structures reported here have been deposited in the PDB under the accession codes 6S8T and 6S8U. Additional data supporting the findings reported in this manuscript are available from the corresponding author on request.

Acknowledgments:

The authors are grateful for the assistance of Ed Lowe and David Staunton (Department of Biochemistry, University of Oxford). We would also like to acknowledge the beamline staff at IO4-1, IO3 and B21 for support during data collection, Anja Jensen for constructs expressing A-type DBL β domains, Natalie Barber for CIDR α domains and Clinton Lau for EPCR. Funding was provided by a Wellcome Trust project grant (087692/Z/08/Z) to MKH and AGC and a Wellcome Investigator Award (101020/Z/13/Z) to MKH.

Author contributions:

F.L., A.G.C and M.K.H. designed research; F.L. and C.S. performed research; F.L. and M.K.H. analysed data; and F.L. and M.K.H. wrote the paper.

The authors declare no conflict of interest.

References:

1. World Health Organisation (2018) World Malaria Report
2. Baruch DI, Gormely JA, Ma C, Howard RJ, Pasloske BL (1996) *Plasmodium falciparum* erythrocyte membrane protein 1 is a parasitized erythrocyte receptor for adherence to CD36, thrombospondin, and intercellular adhesion molecule 1. *Proc Natl Acad Sci U S A* 93(8):3497–3502.
3. Gardner MJ, et al. (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419(6906):498–511.
4. Rask TS, Hansen DA, Theander TG, Pedersen AG, Lavstsen T (2010) *Plasmodium falciparum* erythrocyte membrane protein 1 diversity in seven genomes - divide and conquer. *PLoS Comput Biol* 6(9):e1000933
5. Guizetti J, Scherf A (2013) Silence, activate, poise and switch! Mechanisms of antigenic variation in *Plasmodium falciparum*. *Cell Microbiol* 15(5):718–726.
6. Kraemer SM, Smith JD (2003) Evidence for the importance of genetic structuring to the structural and functional specialization of the *Plasmodium falciparum* var gene family. *Mol Microbiol* 50(5):1527–1538.
7. Lavstsen T, Salanti A, Jensen AT, Arnot DE, Theander TG (2003) Sub-grouping of *Plasmodium falciparum* 3D7 var genes based on sequence analysis of coding and non-coding regions. *Malaria Journal* 2:27
8. Avril M, et al. (2012) A restricted subset of var genes mediates adherence of *Plasmodium falciparum*-infected erythrocytes to brain endothelial cells. *Proc Natl Acad Sci* 109(26):E1782–E1790.
9. Lavstsen T, et al. (2012) *Plasmodium falciparum* erythrocyte membrane protein 1 domain cassettes 8 and 13 are associated with severe malaria in children. *Proc Natl*

- Acad Sci* 109(26):E1791–E1800.
10. Claessens A, et al. (2012) A subset of group A-like var genes encodes the malaria parasite ligands for binding to human brain endothelial cells. *Proc Natl Acad Sci* 109(26):E1772–E1781.
 11. Hviid L, Jensen ATR (2015) PfEMP1 - a parasite protein family of key importance in *Plasmodium falciparum* malaria immunity and pathogenesis. *Adv Parasitol* 88:51–84.
 12. Chesnokov O, Merritt J, Tcherniuk SO, Milman N, Oleinikov A V. (2018) *Plasmodium falciparum* infected erythrocytes can bind to host receptors integrins $\alpha V\beta 3$ and $\alpha V\beta 6$ through DBL $\delta 1_D4$ domain of PFL2665c PfEMP1 protein. *Sci Rep* 8(1) 17871
 13. Ockenhouse CF, Tandon NN, Magowan C, Jamieson GA, Chulay JD (1989) Identification of a platelet membrane glycoprotein as a falciparum malaria sequestration receptor. *Science* 243(4897):1469–1471.
 14. Turner L, et al. (2013) Severe malaria is associated with parasite binding to endothelial protein C receptor. *Nature* 498(7455):502–505.
 15. McCormick CJ, Craig A, Roberts D, Newbold CI, Berendt AR (1997) Intercellular adhesion molecule-1 and CD36 synergize to mediate adherence of *Plasmodium falciparum*-infected erythrocytes to cultured human microvascular endothelial cells. *J Clin Invest* 100(10):2521–2529.
 16. Hsieh FL, et al. (2016) The structural basis for CD36 binding by the malaria parasite. *Nat Commun* 7:12837.
 17. Robinson BA, Welch TL, Smith JD (2003) Widespread functional specialization of *Plasmodium falciparum* erythrocyte membrane protein 1 family members to bind CD36 analysed across a parasite genome. *Mol Microbiol* 47(5):1265–1278.
 18. Lau CKY, et al. (2015) Structural conservation despite huge sequence diversity allows EPCR binding by the PfEMP1 family implicated in severe childhood malaria. *Cell Host Microbe* 17(1):118–129.
 19. Jespersen JS, et al. (2016) *Plasmodium falciparum* var genes expressed in children with severe malaria encode CIDR $\alpha 1$ domains . *EMBO Mol Med* 8(8):839–850.
 20. Bernabeu M, Smith JD (2017) EPCR and Malaria Severity: The Center of a Perfect Storm. *Trends Parasitol* 33(4):295–308.
 21. Mkumbaye SI, et al. (2017) The Severity of *Plasmodium falciparum* Infection Is Associated with Transcript Levels of var Genes Encoding Endothelial Protein C

- Receptor-Binding *P. falciparum* Erythrocyte Membrane Protein 1 . *Infect Immun* 85(4):e00841-16.
22. Lennartz F, et al. (2017) Structure-Guided Identification of a Family of Dual Receptor-Binding PfEMP1 that Is Associated with Cerebral Malaria. *Cell Host Microbe* 21(3):403–414.
 23. Newbold C, et al. (1997) Receptor-specific adhesion and clinical disease in *Plasmodium falciparum*. *Am J Trop Med Hyg* 57(4):389–398.
 24. Rogerson SJ, et al. (1999) Cytoadherence characteristics of *Plasmodium falciparum*-infected erythrocytes from Malawian children with severe and uncomplicated malaria. *Am J Trop Med Hyg* 61(3):467–472.
 25. Ochola LB, et al. (2011) Specific receptor usage in *Plasmodium falciparum* cytoadherence is associated with disease outcome. *PLoS One* 6(3):e14741.
 26. Bengtsson A, et al. (2013) A Novel Domain Cassette Identifies *Plasmodium falciparum* PfEMP1 Proteins Binding ICAM-1 and Is a Target of Cross-Reactive, Adhesion-Inhibitory Antibodies. *J Immunol* 190(1):240–249.
 27. Carrington E, et al. (2018) In silico guided reconstruction and analysis of ICAM-1-binding var genes from *Plasmodium falciparum*. *Sci Rep* 8(1):3282.
 28. Howell DPG, et al. (2008) Mapping a common interaction site used by *Plasmodium falciparum* Duffy binding-like domains to bind diverse host receptors. *Mol Microbiol* 67(1):78–87.
 29. Janes JH, et al. (2011) Investigating the host binding signature on the *Plasmodium falciparum* PfEMP1 protein family. *PLoS Pathog* 7(5):e1002032.
 30. Avril M, Bernabeu M, Benjamin M, Brazier AJ, Smith JD (2016) Interaction between endothelial protein C receptor and intercellular adhesion molecule 1 to mediate binding of *Plasmodium falciparum*-infected erythrocytes to endothelial cells. *MBio* 7(4):e00615-6
 31. Tuikue Ndam N, et al. (2017) Parasites Causing Cerebral Falciparum Malaria Bind Multiple Endothelial Receptors and Express EPCR and ICAM-1-Binding PfEMP1. *J Infect Dis* 215(12):1918–1925.
 32. Casasnovas JM, Stehle T, Liu JH, Wang JH, Springer TA (1998) A dimeric crystal structure for the N-terminal two domains of intercellular adhesion molecule-1. *Proc Natl Acad Sci U S A* 95:4134–4139.

33. Higgins MK, Carrington M (2014) Sequence variation and structural conservation allows development of novel function and immune evasion in parasite surface protein families. *Protein Sci* 23(4):354–365.
34. Remaut H, Waksman G (2006) Protein-protein interaction through β -strand addition. *Trends Biochem Sci* 31(8):436–444.
35. Tse MT, Chakrabarti K, Gray C, Chitnis CE, Craig A (2004) Divergent binding sites on intercellular adhesion molecule-1 (ICAM-1) for variant *Plasmodium falciparum* isolates. *Mol Microbiol* 51(4):1039–1049.
36. Madkhali AM, et al. (2014) An analysis of the binding characteristics of a panel of recently selected ICAM-1 binding *Plasmodium falciparum* patient isolates. *PLoS One* 9(10):e111518.
37. Karlsson R, Fält A (1997) Experimental design for kinetic analysis of protein-protein interactions with surface plasmon resonance biosensors. *J Immunol Methods* 200(1–2):121–133.
38. Juillerat A, et al. (2011) Structure of a *Plasmodium falciparum* PfEMP1 rosetting domain reveals a role for the N-terminal segment in heparin-mediated rosette inhibition. *Proc Natl Acad Sci* 108(13):5243–5248.
39. Dominguez C, Boelens R, Bonvin AMJJ (2003) HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* 125(7):1731–1737.
40. Karaca E, Bonvin AMJJ (2013) On the usefulness of ion-mobility mass spectrometry and SAXS data in scoring docking decoys. *Acta Crystallogr Sect D Biol Crystallogr* 69(5):683–694.
41. Smith JD, Rowe JA, Higgins MK, Lavstsen T (2013) Malaria's deadly grip: Cytoadhesion of *Plasmodium falciparum*-infected erythrocytes. *Cell Microbiol* 15(12):1976–1983.
42. Tessema SK, et al. (2018) Antibodies to Intercellular Adhesion Molecule 1-Binding *Plasmodium falciparum* Erythrocyte Membrane Protein 1-DBL β Are Biomarkers of Protective Immunity to Malaria in a Cohort of Young Children from Papua New Guinea. *Infect Immun* 86(8):e00485.
43. Turner L, Theander TG, Lavstsen T (2018) Immunization with recombinant *Plasmodium falciparum* erythrocyte membrane protein 1 CIDR1 domains induces domain subtype inhibitory antibodies. *Infect Immun* 86(11):e00435-18.

44. Lennartz F, et al. (2015) Mapping the Binding Site of a Cross-Reactive *Plasmodium falciparum* PfEMP1 Monoclonal Antibody Inhibitory of ICAM-1 Binding. *J Immunol* 195(7):3273–3283.
45. Olsen RW, et al. (2018) Natural and vaccine-induced acquisition of cross-reactive IgG-inhibiting ICAM-1-specific binding of a *Plasmodium falciparum* PfEMP1 subtype associated specifically with cerebral malaria. *Infect Immun* 86(4):e00622-17.
46. Oleinikov A V., et al. (2009) High throughput functional assays of the variant antigen PfEMP1 reveal a single domain in the 3D7 *Plasmodium falciparum* genome that binds ICAM1 with high affinity and is targeted by naturally acquired neutralizing antibodies. *PLoS Pathog* 5(4):e1000386.
47. Bachmann A, et al. (2011) Highly co-ordinated var gene expression and switching in clinical *Plasmodium falciparum* isolates from non-immune malaria patients. *Cell Microbiol* 13(9):1397–1409.
48. Cham GKK, et al. (2009) Sequential, Ordered Acquisition of Antibodies to *Plasmodium falciparum* Erythrocyte Membrane Protein 1 Domains. *J Immunol* 183(5):3356–3363.
49. Brown A, et al. (2013) Molecular architecture of a complex between an adhesion protein from the malaria parasite and intracellular adhesion molecule 1. *J Biol Chem* 288(8):5992–6003.
50. Bernabeu M, et al. (2019) Binding heterogeneity of *Plasmodium falciparum* to engineered 3d brain microvessels is mediated by EPCR and ICAM-1. *MBio* 10(3):e00240-19.
51. Gullingsrud J, Saveria T, Amos E, Duffy PE, Oleinikov A V. (2013) Structure-Function-Immunogenicity Studies of PfEMP1 Domain DBL2 β PF11_0521, a Malaria Parasite Ligand for ICAM-1. *PLoS One* 8(4):e61323.
52. Gray C, McCormick C, Turner G, Craig A (2003) ICAM-1 can play a major role in mediating *P. falciparum* adhesion to endothelium under flow. *Mol Biochem Parasitol* 128(2):187–193.
53. Akhouri RR, Goel S, Furusho H, Skoglund U, Wahlgren M (2016) Architecture of Human IgM in Complex with *P. falciparum* Erythrocyte Membrane Protein 1. *Cell Rep* 14(4):723–736.
54. Bengtsson A, et al. (2013) Transfected HEK293 Cells Expressing Functional

- Recombinant Intercellular Adhesion Molecule 1 (ICAM-1) - A Receptor Associated with Severe *Plasmodium falciparum* Malaria. *PLoS One* 8(7):e69999.
55. Aricescu AR, Lu W, Jones EY (2006) A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr Sect D Biol Crystallogr* 62(10):1243–1250.
 56. Winter G, et al. (2018) DIALS: Implementation and evaluation of a new integration package. *Acta Crystallogr Sect D Struct Biol* 74:85–97.
 57. Potterton L, et al. (2018) CCP 4 i 2: The new graphical user interface to the CCP 4 program suite. *Acta Crystallogr Sect D Struct Biol* 74:68–84.
 58. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr Sect D Biol Crystallogr* 67(4):235–242.
 59. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(4):658–674.
 60. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr Sect D Biol Crystallogr* 66(4):486–501.
 61. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr Sect D Biol Crystallogr* 66(2):213–221.
 62. Rambo RP, Tainer JA (2013) Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* 496(7446):477–481.
 63. Franke D, et al. (2017) ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J Appl Crystallogr* 50:1212–1225.
 64. Konarev P V., Volkov V V., Sokolova A V., Koch MHJ, Svergun DI (2003) PRIMUS: A Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr* 36(5):1277–1282.
 65. Svergun DI (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* 25(pt 4):495–503.
 66. Franke D, Svergun DI (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Crystallogr* 42(2):342–346.
 67. Volkov V V., Svergun DI (2003) Uniqueness of ab initio shape determination in small-angle scattering. *Journal of Applied Crystallography*, pp 860–864.
 68. Svergun DI (1999) Restoring low resolution structure of biological macromolecules

- from solution scattering using simulated annealing. *Biophys J* 76(6):2879–2886.
69. Pettersen EF, et al. (2004) UCSF Chimera - A visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612.
 70. Waterhouse A, et al. (2018) SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res* 46(W1):W296–W303.
 71. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5):1792–1797.
 72. Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol* 18(5):691–699.
 73. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 33(7):1870–1874.
 74. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: A sequence logo generator. *Genome Res* 14(6):1188–1190.
 75. Chattopadhyay R, Taneja T, Chakrabarti K, Pillai CR, Chitnis CE (2004) Molecular analysis of the cytoadherence phenotype of a *Plasmodium falciparum* field isolate that binds intercellular adhesion molecule -1. *Mol Biochem Parasitol* 133(2):255–265.

Figures and Tables:

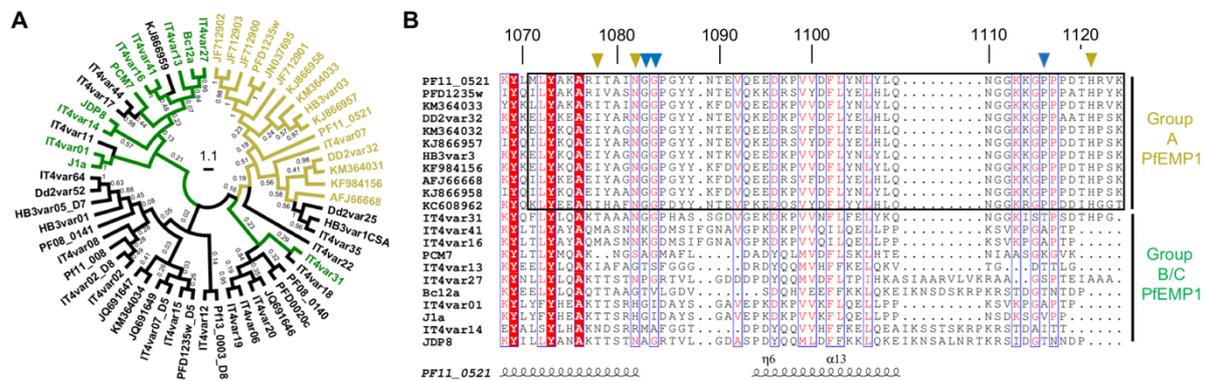


Figure 1 – Group A and group BC PfEMP1 do not share a conserved ICAM-1 binding site

(A) Maximum likelihood tree based on 59 DBL β domain sequences. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Numbers on the branches show bootstrap values. DBL β domains from A-type ICAM-1 binding PfEMP1 are highlighted in yellow, DBL β domains from BC-type ICAM-1 binding PfEMP1 in green and DBL β domains shown not to bind to ICAM-1 are in black (see also Table S1). **(B)** Multiple sequence alignment of DBL β domains from ICAM-1 binding PfEMP1. Numbers show positions in the PF11_0521 DBL β domain. The box highlights the A-type ICAM-1 binding site. Yellow triangles are residues critical for direct interaction of A-type PfEMP1 with ICAM-1. Blue triangles are residues important for the conformation of the A-type ICAM-1 binding site.

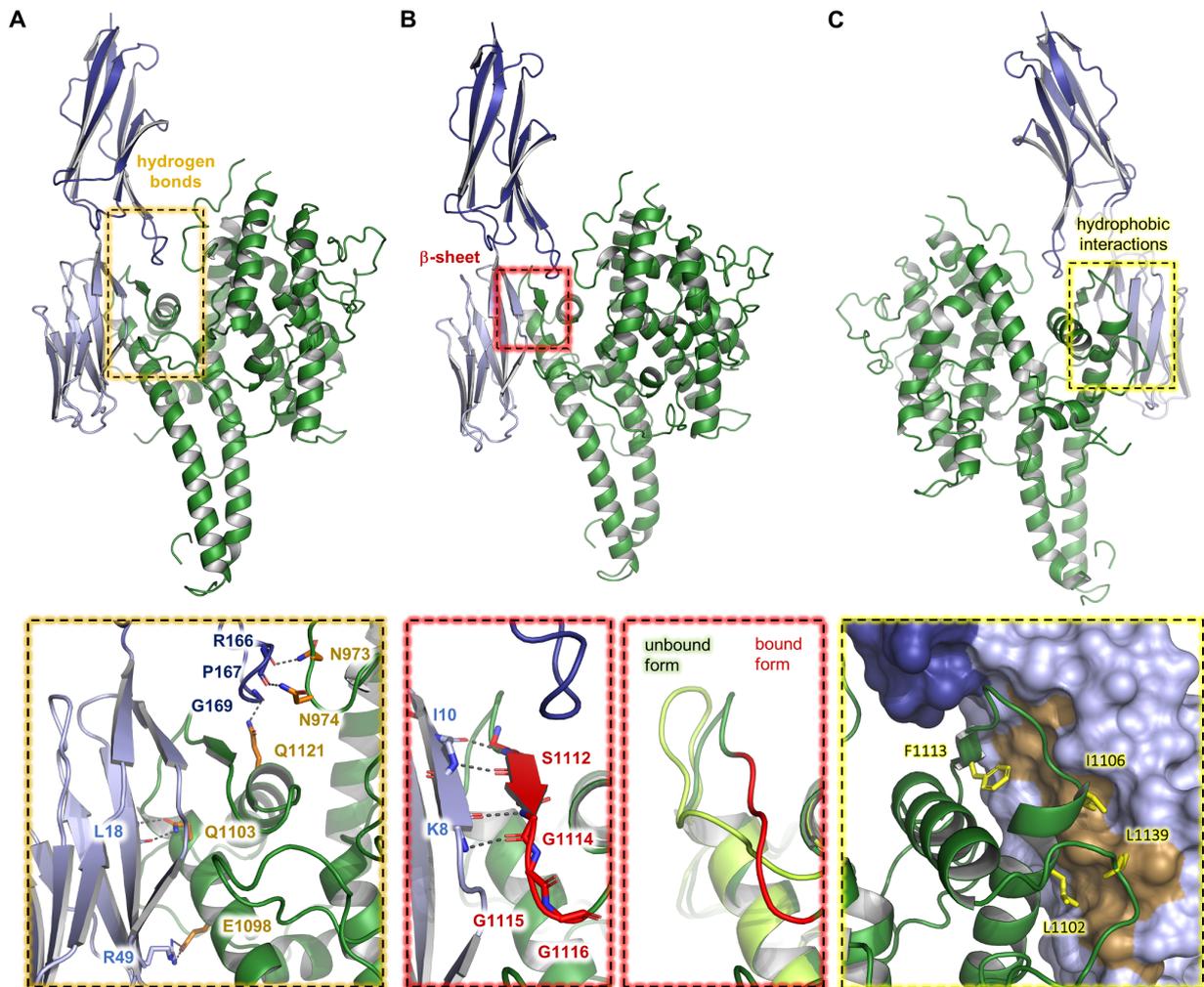


Figure 2. The structural basis for ICAM-1 binding by group B PfEMP1

Front views of the DBL β domain of IT4var13 (green) bound to ICAM-1^{D1D2} (D1 light blue, D2 dark blue). Dashed boxes highlight the sites that contact ICAM-1 through **(A)** side-chain mediated hydrogen bonds or **(B)** a β -sheet augmentation. A third dashed box compares a region of the ICAM-1-bound and unbound conformations of the IT4var13 DBL β domain. **(C)** Back view of the IT4var13 DBL β -ICAM-1^{D1D2} complex. The dashed box highlights the site of hydrophobic contacts between the DBL β domain and ICAM-1^{D1D2} and the residues in the DBL β domain that contact a hydrophobic patch on the surface of ICAM-1^{D1} (dark yellow).

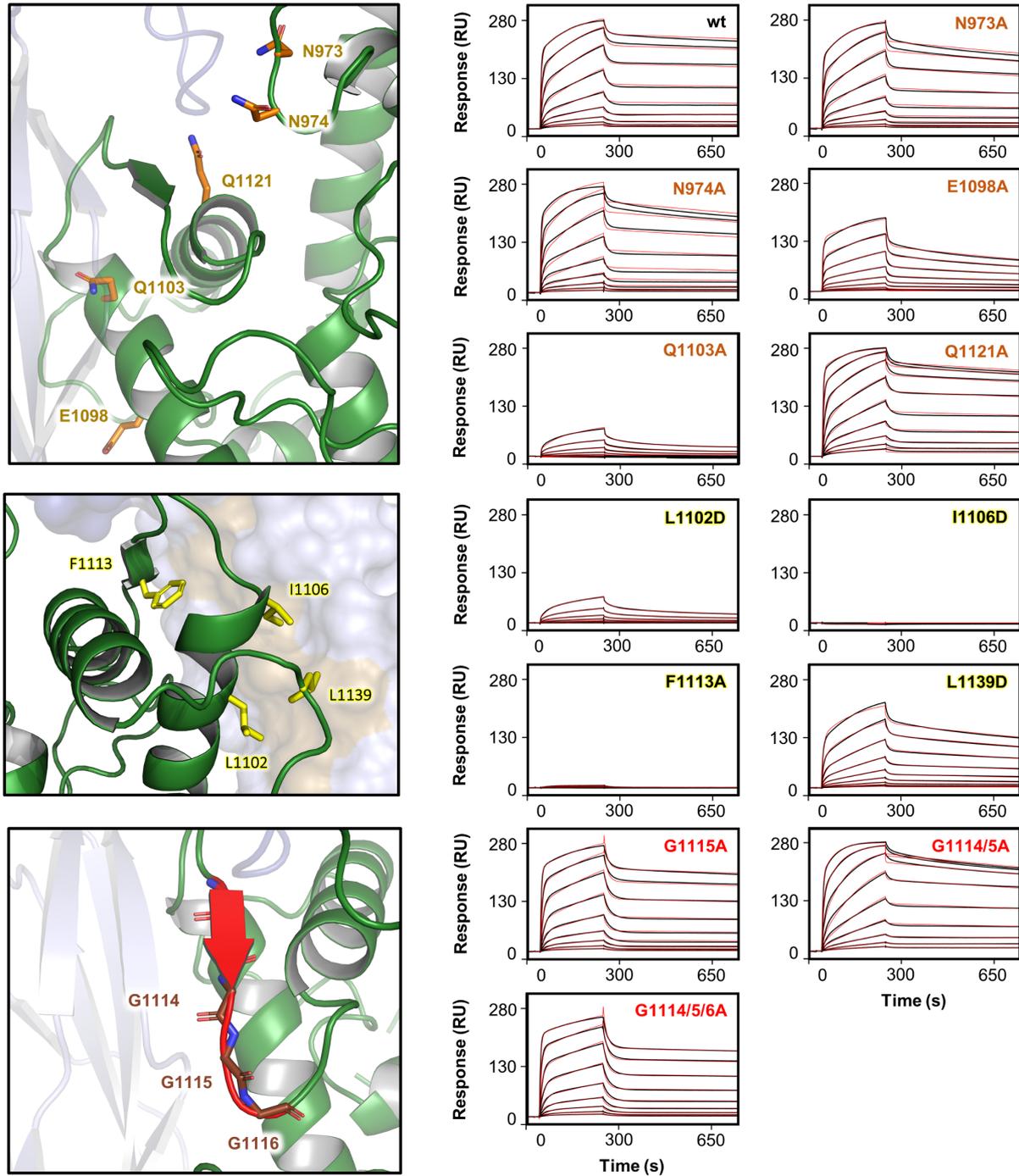


Figure 3 – Mutational analysis of the ICAM-1 binding site from group B PfEMP1

The different sites of contact between the IT4var13 DBL β domain and ICAM-1^{D1D2} and the residues involved in the binding site are shown. For SPR measurements, two-fold dilution series from 250nM to 0.9nM of wild type or mutant IT4var13 DBL β were injected over ICAM-1^{D1D5}-Fc immobilized on a Protein A sensor chip. Sensorgrams show the data (black lines) and the fit of a two-state reaction model (red lines).

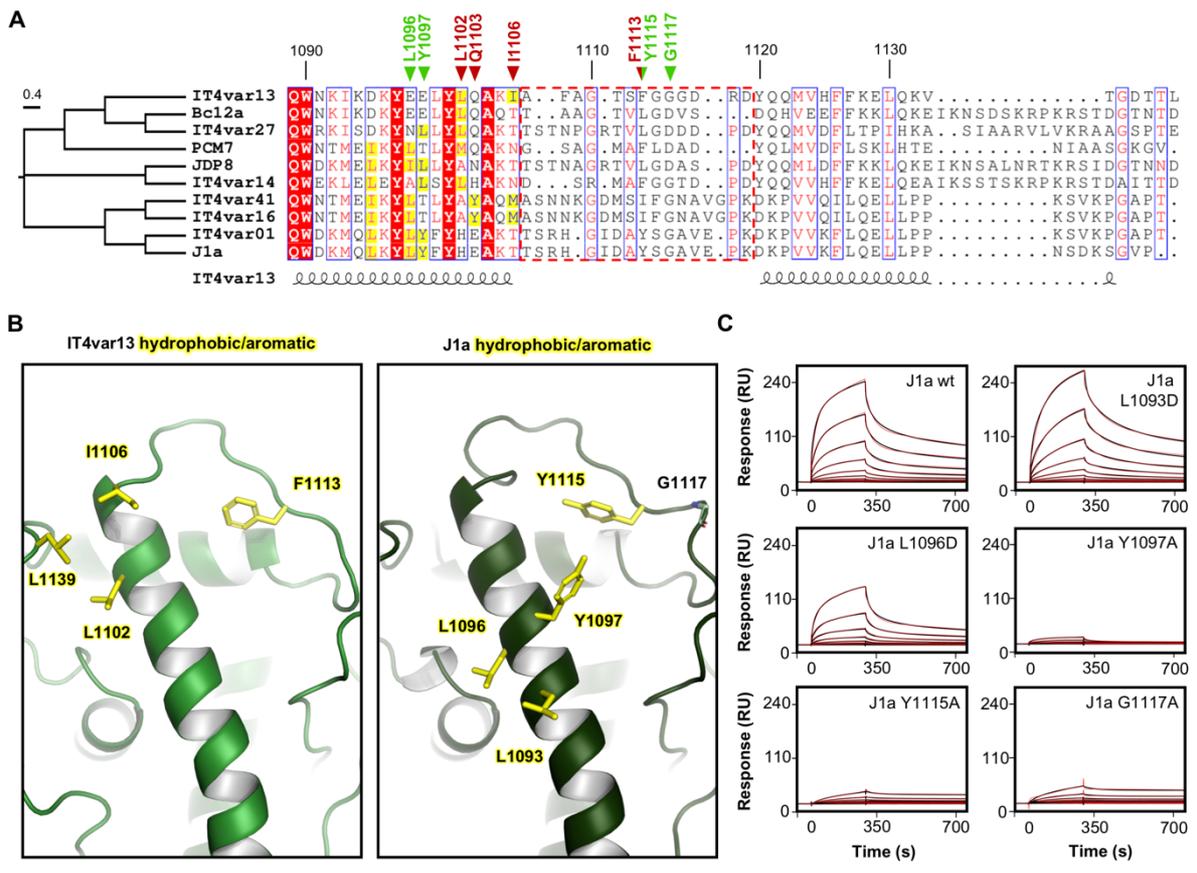


Figure 4 – Limited conservation of the ICAM-1 binding site among group BC PfEMP1

(A) Multiple sequence alignment of DBLβ domains known to bind ICAM-1. Numbers indicate positions in the IT4var13 DBLβ domain. Residues critical for the IT4var13 DBLβ-ICAM-1 interaction are marked with red triangles while those critical for ICAM-1 binding by J1a DBLβ are marked with green triangles. A half green and half red triangle marks a residue important for both IT4var13 and J1a DBLβ to bind ICAM-1. The flexible loop is indicated with a dashed red box. Surface-exposed hydrophobic and aromatic residues along the helix that forms part of the ICAM-1 binding site are highlighted in yellow. **(B)** Positions of hydrophobic and aromatic residues in the ICAM-1 binding site of IT4var13 DBLβ and equivalent positions in a homology model of the J1a DBLβ domain. **(C)** Binding of J1a DBLβ domain and its mutants to ICAM-1. Wild type and mutants were injected in a two-fold dilution series from 500nM to 0.9 nM over ICAM-1^{D1D5}-Fc immobilized on a Protein A sensor chip. Sensorgrams show the data (black lines) and the fit of a two-state reaction model (red lines).

(red) DBL β are indicated. The inset shows a magnification of the ICAM-1 binding sites of the individual domains. **(D)** Overlay of the ICAM-1 bound complexes of IT4var13 DBL β and PF11_0521 DBL β , with the DBL β domains superimposed. The C-terminus of ICAM-1^{D1D2} is indicated. **(E)** Models for complexes of three DBL β domains from A-type PfEMP1 or BC-type PfEMP1 bound to ICAM-1, superimposed on the DBL β domain. The models were derived from docking with HADDOCK, followed by filtering using SEC-SAXS data. **(F)** Schematic representation of the overall binding angles between ICAM-1 and DBL β domains from A-type (red) and BC-type (blue) PfEMP1.

Table 1 – Crystallographic data collection and refinement statistics

	IT4var13 DBL β	IT4var13 DBL β -ICAM-1 ^{D1D2}
Data Collection		
Space group	P12 ₁ 1	P6 ₅ 22
Unit cell parameters		
a, b, c [Å]	67.3, 43.4, 87.1	142.4, 142.4, 224.2
α , β , γ [°]	90.0, 103.6, 90.0	90.0, 90.0, 120.0
Resolution [Å]	32.61 - 2.17 (2.21 - 2.17)	112.11 - 3.67 (3.73 - 3.67)
R _{merge}	0.12 (1.55)	0.24 (1.88)
R _{pim}	0.07 (0.92)	0.08 (0.63)
I/sigI	5.30 (1.56)	6.50 (1.02)
CC _{1/2}	0.99 (0.30)	0.98 (0.55)
Completeness [%]	99.81 (96.85)	100 (100)
Redundancy	4.34 (4.35)	9.83 (9.34)
Wilson B Factor (Å ²)	40.58	124.80
No. of unique reflections	26322 (1229)	15394 (758)
Refinement		
R _{work} /R _{free}	19.96 / 22.85	24.03 / 28.65
No. atoms		
Protein	3600	4899
Ligand	-	56
Water	122	-
B-Factors (Å ²)		
Protein	62.85	160.02
Ligand	-	215.66
Water	48.81	-
R.m.s. deviations		
Bond lengths [Å]	0.003	0.003
Bond angles [°]	0.457	0.701

Values in parentheses are for the highest resolution shell.

Supplementary Information

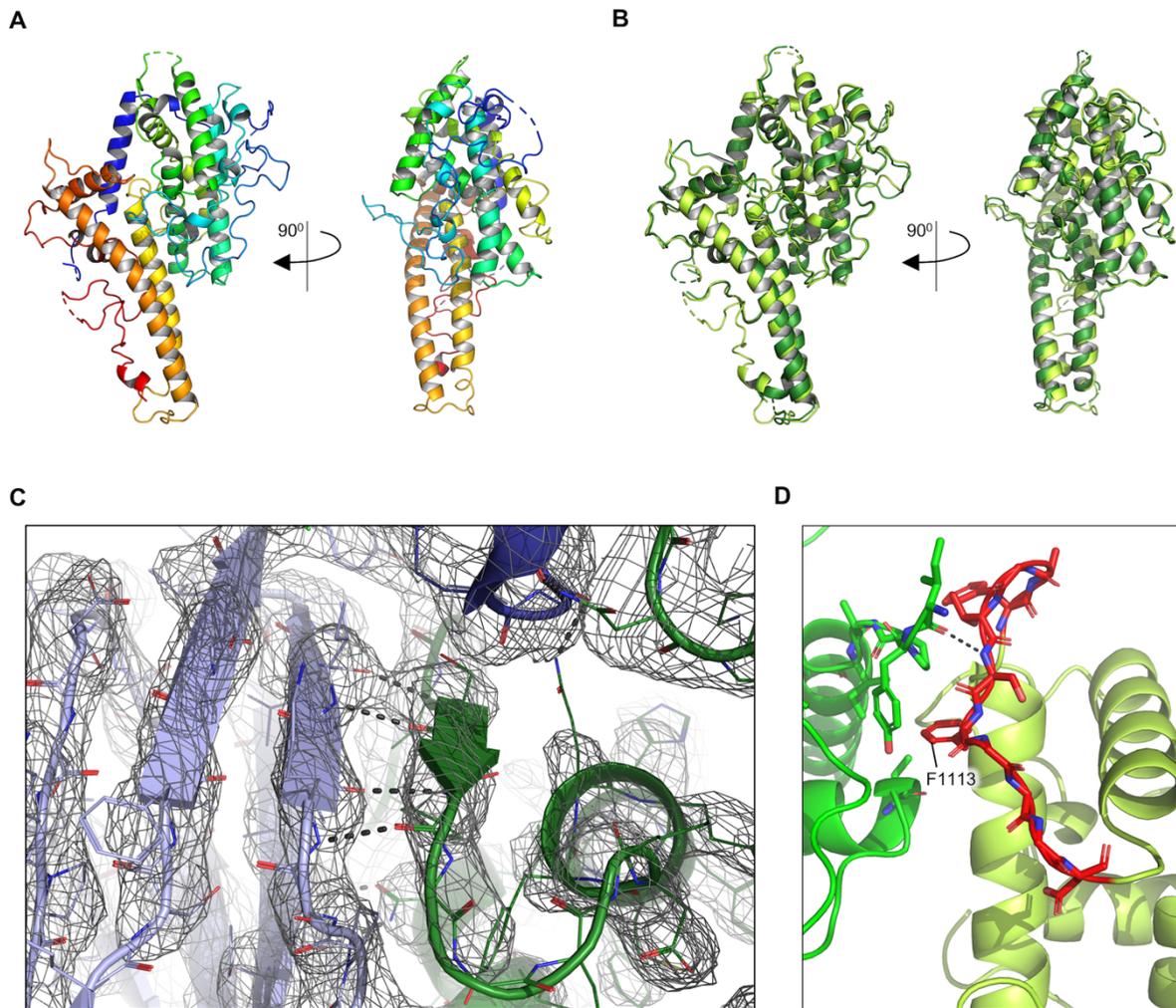


Fig. S1 – Comparison of the structure of IT4var13 DBL β in the ICAM-1^{D1D2} bound and unbound forms

(A) Structure of the IT4var13 DBL β domain in its ICAM-1 bound form. The domain is color coded from the N-terminus (blue) to the C-terminus (red). **(B)** Superposition of the structure of the IT4var13 DBL β in its unbound (bright green) and ICAM-1 bound (dark green) forms. **(C)** Sample electron density (feature-enhanced 2FoFc map) for the ICAM-1 binding site in the IT4var13 DBL β -ICAM-1^{D1D2} complex, contoured at a sigma level of 1.0. Dashed lines indicate hydrogen bonds. **(D)** Crystal contacts around the flexible loop of the ICAM-1 binding site (red) between symmetry mates (green and bright green) in the crystals of the unbound IT4var13 DBL β domain. Dashed lines indicate hydrogen bonds. For orientation, the residue F1113 is indicated.

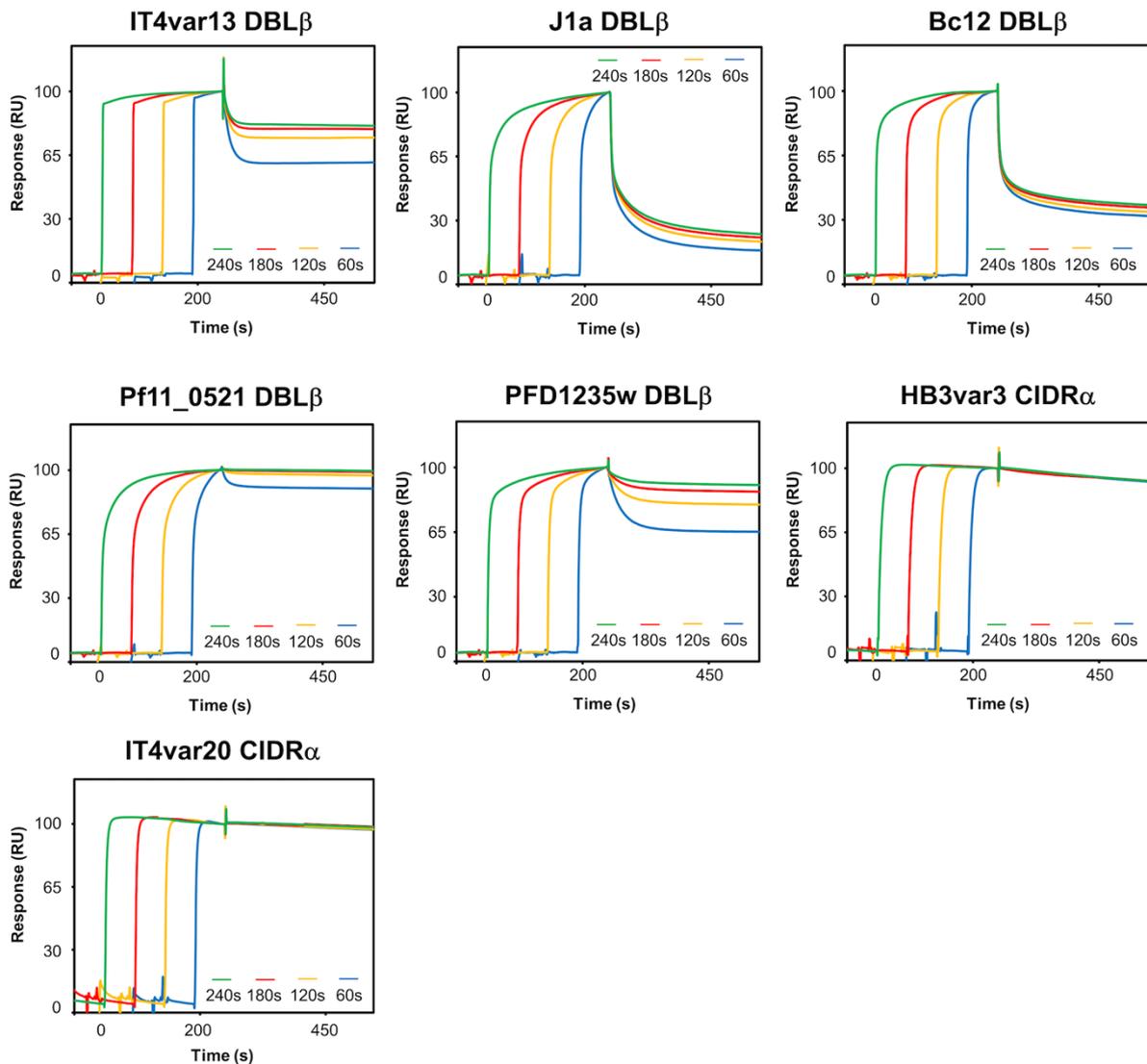


Figure S2 – SPR analysis of potential conformational changes upon receptor binding for DBL β and CIDR α domains

Domains were injected over their respective receptors at a fixed concentration of $3\mu\text{M}$. For each injection, the association time was varied in 60s intervals between 60s and 240s. The data were then superimposed using the end of the injection as a fixed time point and normalized by setting the signal at the start of the injection to 0 and at the end of the injection to 100, to monitor how the dissociation rate changes with the length of association time. DBL β domains were injected over ICAM-1^{D1D5}-Fc immobilized on a CM5 chip pre-coupled with Protein A. CIDR α domains were injected over biotinylated EPCR immobilized on a CAP sensor chip.

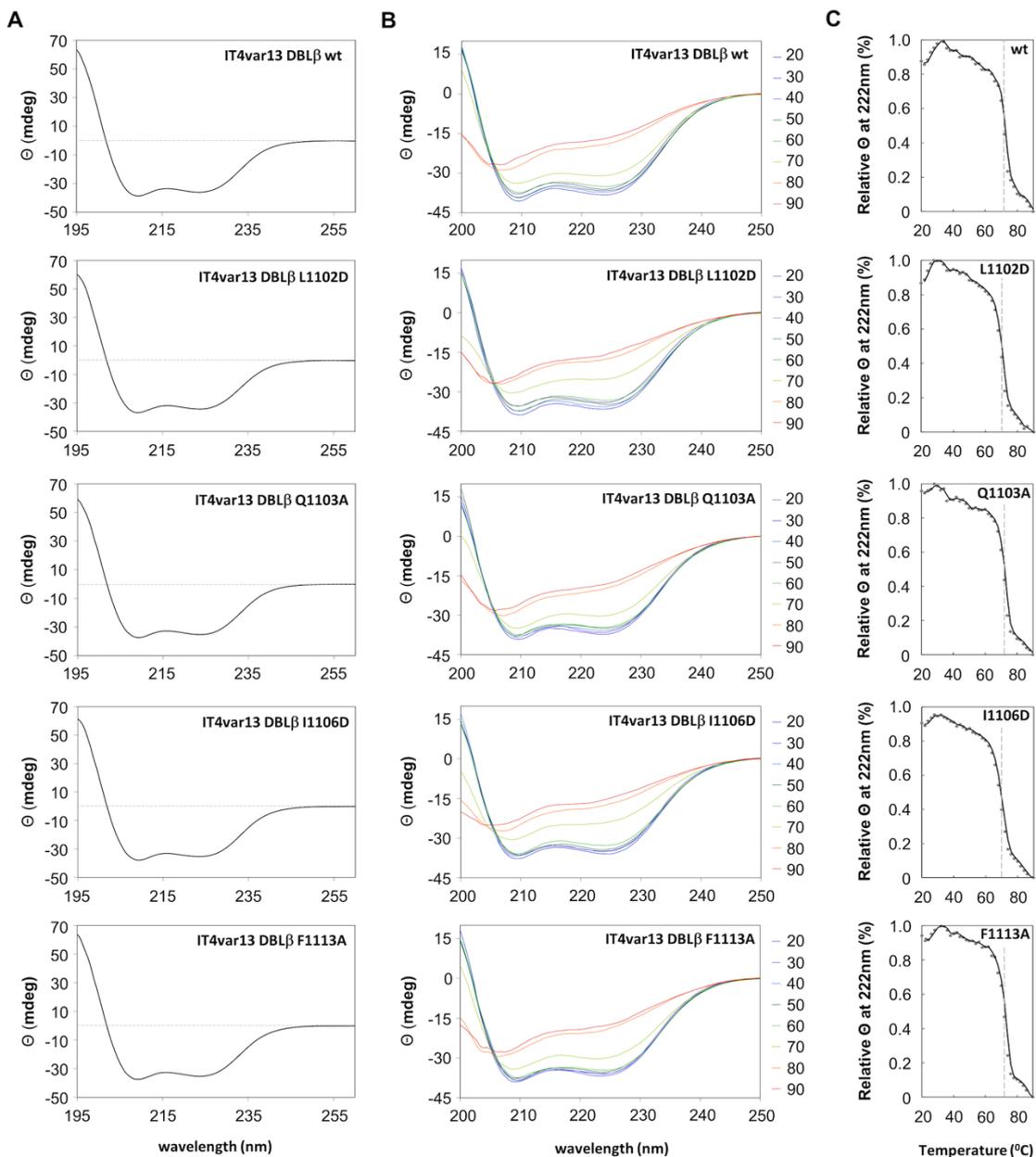


Figure S3 – CD spectroscopy and melting curves for IT4var13 DBL β wild type and mutants

(A) Secondary structure analysis of the IT4var13 DBL β wild type (wt) domain and mutants. CD spectra were recorded at 20°C between wavelengths of 195 nm and 260 nm. For each sample, four measurements were averaged and corrected for buffer absorption. **(B)** Thermal melt analysis of IT4var13 DBL β wild type and mutant domains. CD spectra were recorded between 200 nm and 250 nm and after each measurement the temperature was raised by 0.5°C. Shown are spectra collected at 10°C intervals between 20°C and 90°C. **(C)** Thermal denaturation curve for IT4var13 DBL β wild type and mutants. The ellipticity at a fixed wavelength of 222 nm was measured between 20°C and 90°C. After each measurement, the temperature was increased by 0.5°C. A dashed line indicates the melting point.

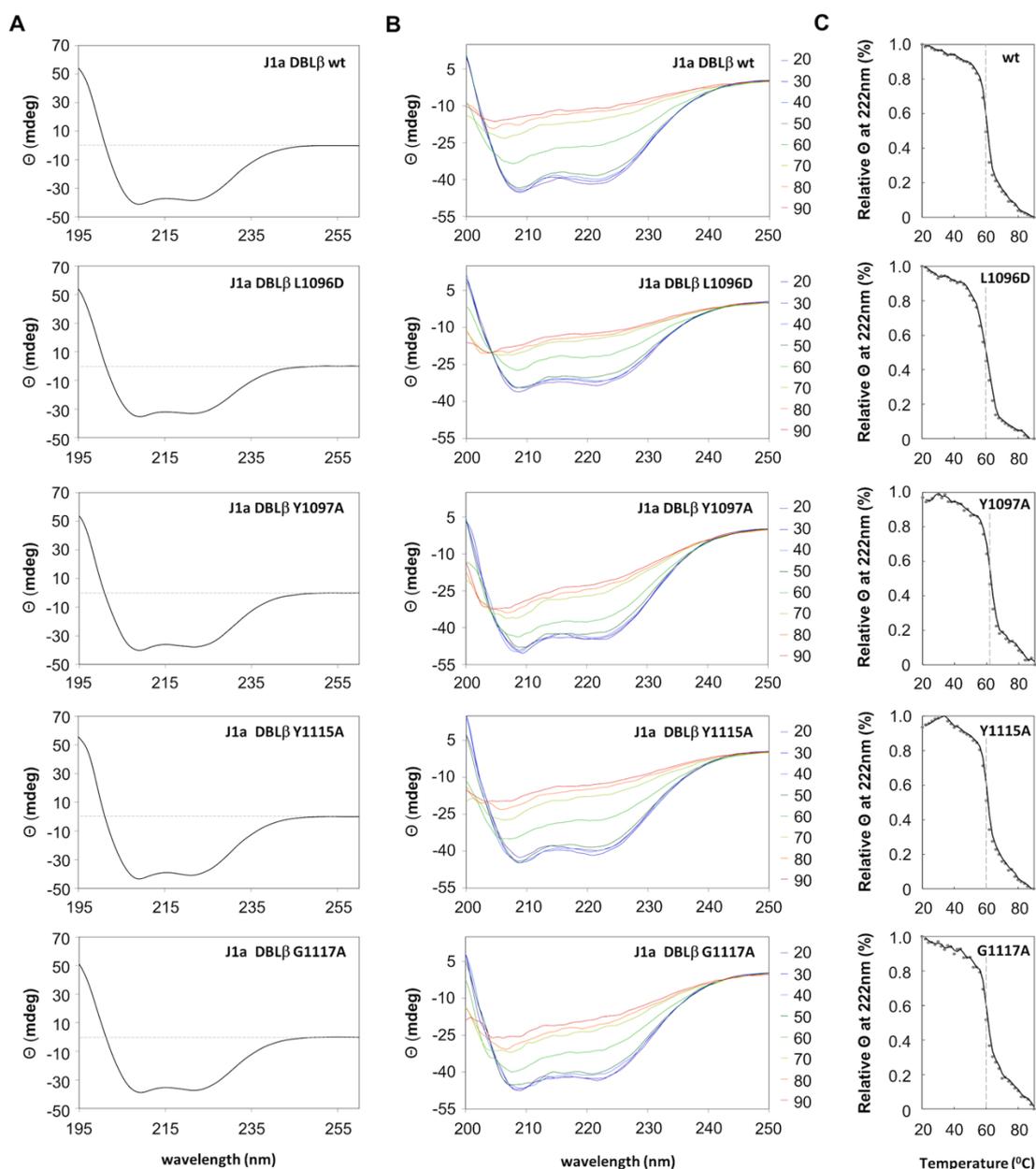


Figure S4 – CD spectroscopy and melting curves for J1a DBL β wild type and mutants

(A) Secondary structure analysis of the J1a DBL β wild type (wt) domain and mutants. CD spectra were recorded at 20°C between wavelengths of 195 nm and 260 nm. For each sample, four measurements were averaged and corrected for buffer absorption. **(B)** Thermal melt analysis of J1a DBL β wild type and mutant domains. CD spectra were recorded between 200 nm and 250 nm and after each measurement the temperature was raised by 0.5°C. Shown are spectra for 10°C intervals between 20°C and 90°C. **(C)** Thermal denaturation curve for J1a DBL β wild type and mutants. The ellipticity at a fixed wavelength of 222 nm was measured between 20°C and 90°C. After each measurement, the temperature was increased by 0.5°C. A dashed line indicates the melting point.

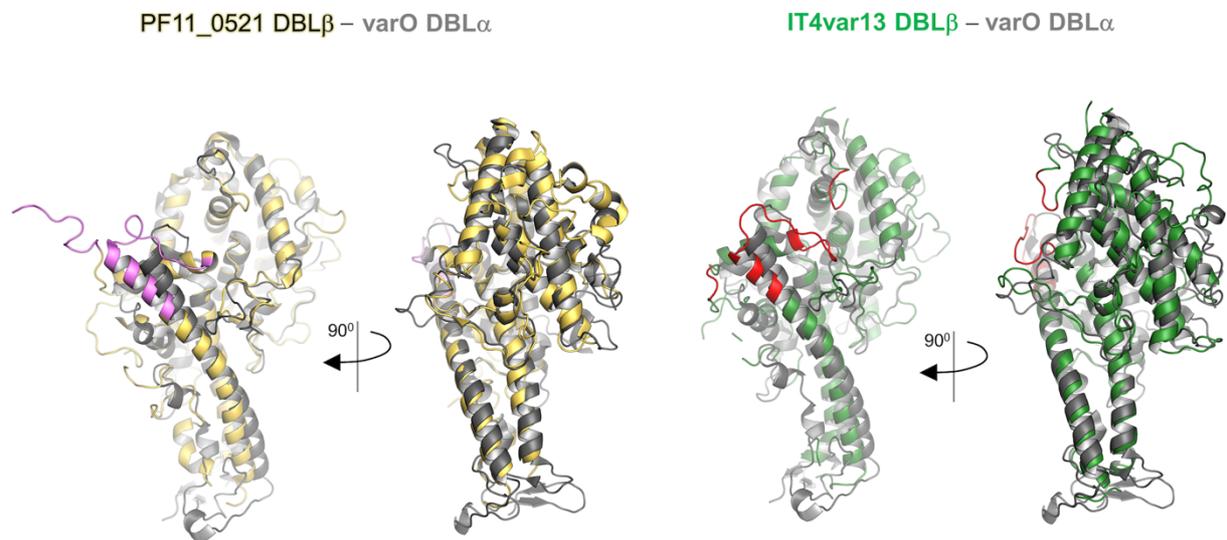


Figure S5 - Superposition of the DBL β domains from IT4var13 and PF11_0521 PfEMP1 onto the DBL α domain from the varO PfEMP1.

IT4var13 DBL β is green and the ICAM-1 binding site of this domain is red. PF11_0521 DBL β is yellow and the ICAM-1 binding site of this domain is pink. The varO DBL α domain (PDB 2XU0) (38) is grey. The superposition was made using PyMol.

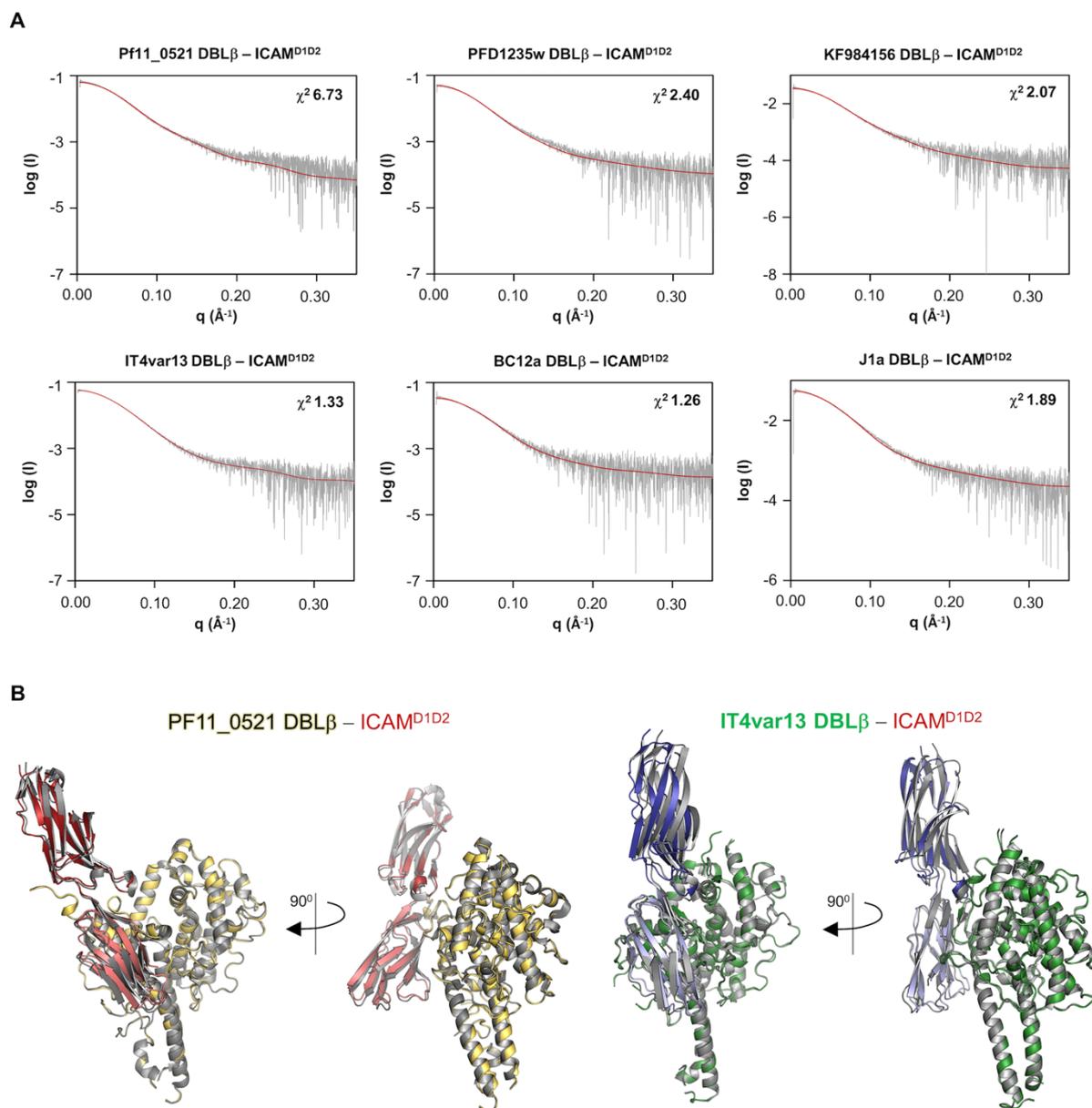


Figure S6. Crysol fits for DBL β -ICAM-1^{D1D2} models generated by HADDOCK and comparison with the crystal structures

(A) Crysol fits between SAXS data and the top models for DBL β -ICAM-1^{D1D2} complexes selected after SAXS filtering. The models for **(B)** the PF11_0521 DBL β -ICAM-1^{D1D2} (DBL β : dark yellow; ICAM-1^{D1D2}: red) and **(C)** the IT4var13 DBL β - ICAM-1^{D1D2} (DBL β : green; ICAM-1^{D1D2}: blue) complexes derived from HADDOCK after SAXS-filtering are superimposed over the crystal structures of the respective complexes (grey).

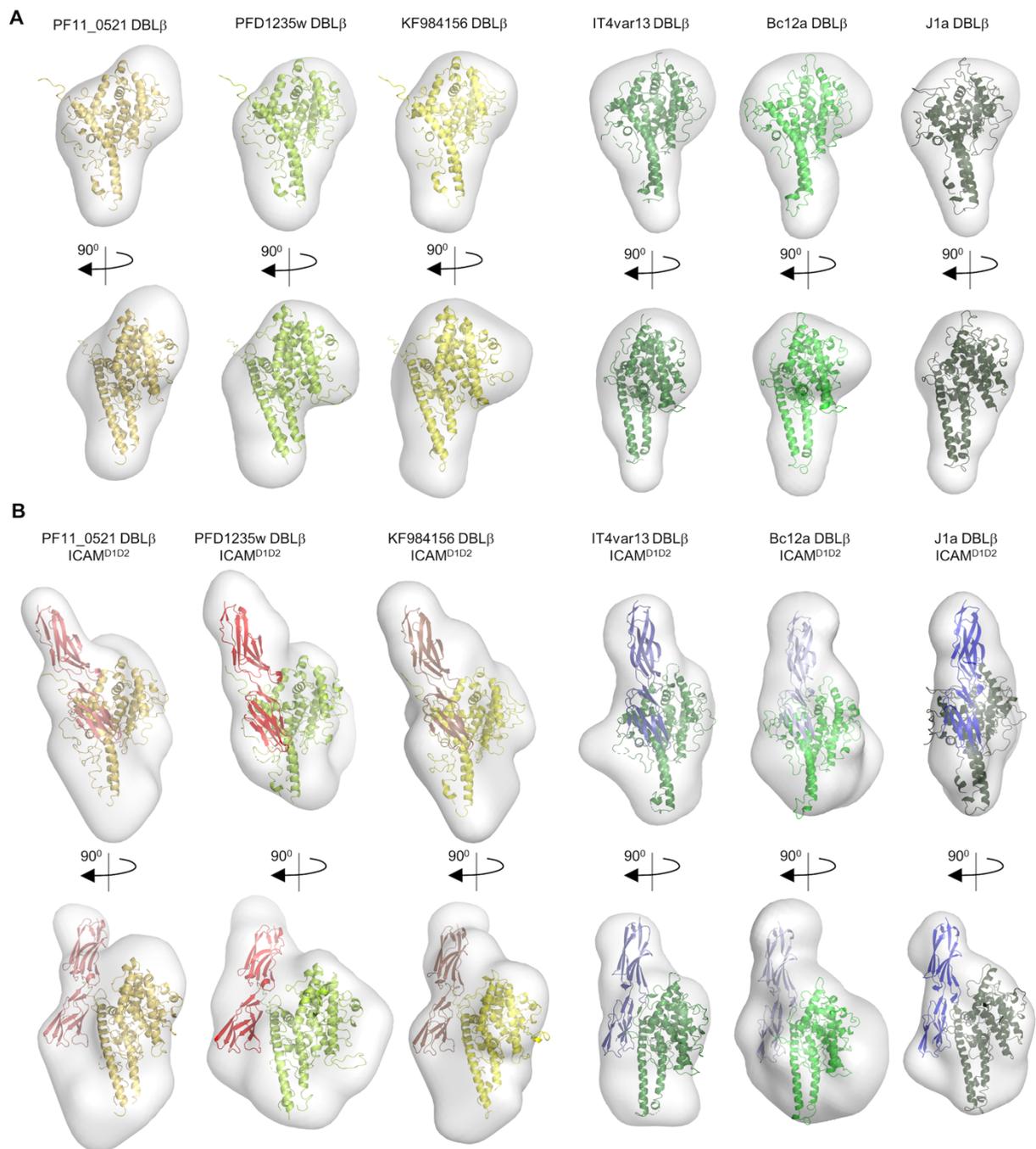


Figure S7 – SAXS analysis of DBL β -ICAM-1^{D1D2} complexes

Ab initio envelopes calculated from SAXS data for DBL β domains from group A and group BC PfEMP1, either **(A)** alone or **(B)** in complex with ICAM-1^{D1D2}. The models were generated by SwissModel for the DBL β domains alone or by HADDOCK for the DBL β domains bound to ICAM-1^{D1D2} and were docked into the envelopes using Chimera.

Table S1 Sequences used for phylogenetic analyses

PfEMP1	GenBank accession number	ICAM-1 binding phenotype shown in
IT4var01	AAO67411	(28, 29)
IT4var02	AAQ73925	(28, 29)
IT4var02_D8	AAQ73925	(29)
IT4var06	ABM88781	(28, 29)
IT4var07	ABM88782	(30)
IT4var07_D5	ABM88782	(28, 29)
IT4var08	ABM88783	(28, 29)
IT4var11	ETW15085	(29)
IT4var12	AAB06961	(28, 29)
IT4var13	ABM88750	(28, 29, 49)
IT4var14	AAD03351	(28, 29)
IT4var15	AAA75398	(28, 29)
IT4var16	AAS89259	(28, 29, 49)
IT4var17	ABM88751	(28, 29)
IT4var18	ABM88752	(28, 29)
IT4var19	ABM88753	(28, 29)
IT4var20	AAA75397	(28, 29)
IT4var22	ABM88754	(28, 29)
IT4var27	ABM88759	(28, 29, 49)
IT4var31	AAF18980	(28, 29, 49)
IT4var35	ADK78857	(28, 29)
IT4var41	ABM88768	(28, 29, 49)
IT4var44	ABM88769	(29)
IT4var64	ABM88780	(29)
Bc12a	AVI24062	(27)
J1a	AVI24063	(27)
PCM7	AVI24066	(27)
JDP8	AAK49742	(75)
PF08_0140	XP_001349512	(22)

PF08_0141	XP_001349513	(22)
PF11_0521	XP_001348176	(22, 46)
PF11_008	XP_001347692	(22, 26)
PF13_0003_D8	XP_001349740	(22)
PFD1235w	XP_002808895	(22, 26)
PFD1235w_D5	XP_002808895	(22, 26)
PFD0020c	EWC90558	(22)
HB3var01	KOB61483	(22)
HB3var03	KOB63865	(22)
HB3var1CSA	KOB61785	(22)
HB3var05_D7	KOB63865	(22)
Dd2var25	KOB84767	(22)
DD2var32	KOB85388	(22, 26)
Dd2var52	ABM88780	(22)
JF712902	AEQ26010	(22, 26)
JF712903	AEA86277	(22, 26)
JF712900	JF712900	(22, 26)
JF712901	JF712901	(22, 26)
KF984156	AIX97102	(22)
KM364031	AIY25809	(22)
KM364034	AIY25811	(22)
KM364033	AIY25810	(22)
KJ866957	AIX97164	(22)
KJ866958	AIX97165	(22)
KJ866959	AIX97166	(22)
JQ691646	AFJ66676	(22)
JQ691647	AFJ66677	(22)
JQ691649	AFJ66679	(22)
AFJ66668	AFJ66668	(22)
JN037695	AEI26313	(22)

Table S2 – List of contacts between IT4var13DBL β and ICAM-1^{D1D2}

Colors correspond to those used in Fig. 2 of the main text.

IT4var13 DBL β		ICAM-1 ^{D1D2}		Type of Interaction
Residue	Group	Residue	Group	
Asn973	Sidechain NH ₂	Arg166	Backbone O	Hydrogen bond
Asn974	Sidechain NH ₂	Arg167	Backbone O	Hydrogen bond
Glu1098	Sidechain O	Arg49	Sidechain NH ₂	Hydrogen bond
Leu1102	Sidechain	Leu44	Sidechain	Hydrophobic
Leu1102	Sidechain	Val51	Sidechain	Hydrophobic
Gln1103	Sidechain O	Leu18	Backbone N	Hydrogen bond
Gln1103	Sidechain NH ₂	Leu18	Backbone O	Hydrogen bond
Ile1106	Sidechain	Leu18	Sidechain	Hydrophobic
Ser1112	Backbone N	Ile10	Backbone O	Hydrogen bond
Ser1112	Backbone O	Ile10	Backbone N	Hydrogen bond
Phe1113	Sidechain	Val9	Sidechain	Hydrophobic
Phe1113	Sidechain	Leu11	Sidechain	Hydrophobic
Phe1113	Sidechain	Val17	Sidechain	Hydrophobic
Gly1114	Backbone N	Lys8	Backbone O	Hydrogen bond
Gly1114	Backbone O	Lys8	Backbone N	Hydrogen bond
Gln1121	Sidechain NH ₂	Gly169	Backbone O	Hydrogen bond
Leu1139	Sidechain	Leu42	Sidechain	Hydrophobic

Table S3 – Kinetic parameters derived from SPR for binding of ICAM-1^{D1D5} to IT4var13 DBL β wild type and mutants

Interaction	k_{a1} (x10⁵ M⁻¹ s⁻¹)	k_{d1} (10⁻² s⁻¹)	k_{a2} (x10⁻² s⁻¹)	k_{d2} (10⁻⁴ s⁻¹)	K_D (nM)
ICAM ^{D1D5} - IT4var13 DBL β wt	6.47	7.76	1.13	1.17	1.22
ICAM ^{D1D5} - IT4var13 DBL β N973A	36.40	60.80	1.66	3.24	3.22
ICAM ^{D1D5} - IT4var13 DBL β N974A	64.00	99.22	1.53	3.02	2.99
ICAM ^{D1D5} - IT4var13 DBL β E1098A	4.26	9.37	1.21	7.82	13.40
ICAM ^{D1D5} - IT4var13 DBL β L1102D	0.21	2.46	0.54	10.91	196.00
ICAM ^{D1D5} - IT4var13 DBL β Q1103A	0.09	1.47	0.43	8.98	259.00
ICAM ^{D1D5} - IT4var13 DBL β I1106D	0.80	87.74	0.28	452.40	10300.00
ICAM ^{D1D5} - IT4var13 DBL β F1113A	0.05	3.71	0.11	10.31	3620.00
ICAM ^{D1D5} - IT4var13 DBL β Q1121A	13.60	8.49	1.49	2.11	0.86
ICAM ^{D1D5} - IT4var13 DBL β L1139D	3.19	6.07	1.18	5.89	9.06
ICAM ^{D1D5} - IT4var13 DBL β G1115A	8.28	8.37	0.56	1.25	2.21
ICAM ^{D1D5} - IT4var13 DBL β G1114/5A	630.00	685.00	3.07	3.12	1.09
ICAM ^{D1D5} - IT4var13 DBL β G1114/5/6A	7.96	7.54	0.43	0.85	1.86

Table S4 - Kinetic parameters derived from SPR for binding of ICAM-1^{D1D5} to J1a DBL β wild type and mutants

Interaction	k_a1 ($\times 10^4 \text{ M}^{-1} \text{ s}^{-1}$)	k_d1 (10^{-2} s^{-1})	k_a2 ($\times 10^{-3} \text{ s}^{-1}$)	k_d2 (10^{-4} s^{-1})	K_D (nM)
ICAM ^{D1D5} - J1a DBL β wt	1.75	2.54	4.26	9.40	262.00
ICAM ^{D1D5} - J1a DBL β L1093D	1.07	1.07	3.54	10.64	232.00
ICAM ^{D1D5} - J1a DBL β L1096D	0.74	2.55	3.20	10.81	866.00
ICAM ^{D1D5} - J1a DBL β Y1097A	0.08	3.45	3.46	4.74	5140.00
ICAM ^{D1D5} - J1a DBL β Y1115A	0.16	3.73	7.93	1.71	498.00
ICAM ^{D1D5} - J1a DBL β G1117A	4.27	44.15	4.48	1.51	336.00

Table S5 - Parameters derived from solution scattering data

	R_g (nm)	D_{max} (nm)	V_{porod} (nm ³)	Mr_{exp} (kDa)	Mr_{app} (kDa)
DBLβ domain alone					
PF11_0521 DBL β	27.30	94.12	84.0	49.80	57.20
PFD1235w DBL β	26.80	86.43	85.0	50.00	56.90
KF984156 DBL β	28.00	88.07	86.9	51.12	57.90
IT4var13 DBL β	26.00	89.35	85.8	50.47	55.60
Bc12a DBL β	26.00	88.64	78.6	46.26	54.60
J1a DBL β	26.40	85.86	79.0	47.00	53.60
DBLβ-ICAM^{D1D2}					
PF11_0521 DBL β – ICAM-1 ^{D1D2}	35.11	126.10	140.0	89.4	82.00
PFD1235w DBL β – ICAM-1 ^{D1D2}	35.21	128.07	132.0	90.0	82.00
KF984156 DBL β – ICAM-1 ^{D1D2}	35.33	127.74	139.0	94.2	82.00
IT4var13 DBL β – ICAM-1 ^{D1D2}	32.40	110.49	133.0	86.3	82.00
Bc12a DBL β – ICAM-1 ^{D1D2}	33.74	113.41	115.0	74.9	80.00
J1a DBL β – ICAM-1 ^{D1D2}	31.08	109.00	98.8	68.6	80.00

The radius of gyration R_g was determined using AutoRg and the maximum particle diameter D^{max} was calculated with GNOM. The Porod volume V_{porod} was calculated with PRIMUS. The apparent molecular mass Mr_{app} was calculated using the from the size and shape function in version 2.8.3 of the ATSAS suite. Mr_{exp} is the theoretical molecular mass.